Engineering Light-Energy Conversion into Nonphotosynthetic Hosts

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Ilya Tikh

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Claudia Schmidt-Dannert

December 2013

Acknowledgements

First, I must thank Claudia for her guidance and patience over the years it took me figure out how to be a good scientist. Thank you for being a supportive mentor and not discarding some of my crazier ideas.

I want to also thank all of the members of the Schmidt-Dannert lab, past and present, who have helped and supported me over the years. I am especially grateful to Ethan and Maureen, as both of you have shown tremendous patience with me and ultimately helped me to become the scientist I am today.

I want to thank everyone in BTI for your thoughtful discussions and not minding occasional reagent borrowing.

I am grateful to my chair, Dr. Michael Sadowsky and my committee members; Dr. Burckhard Seelig, Dr. Jeff Gralnick and Dr. Do-Hyung Kim for their guidance over the years.

I was fortunate to start graduate school with some amazing peers who have become family over the years. Thank you for being there, good times and bad, Val, Lauren, Zach, Joel, Aaron and Eu Han.

I want to thank Joni being patient with me and my continuisly evolving graduation timelines.

Finally, I want to thank my family for their continuous and unwavering support over the last six years and for not asking me when I would graduate, too frequently.

Dedication

To my grandparents

Abstract

Over billions of years photosynthetic organisms have refined the molecular machinery required for the capture and conversion of light into chemical energy. To date, much research has been devoted into harnessing this unique trait from photosynthetic organisms and utilizing them for ecologically clean production of valuable resources, such as alternatives to fossil fuels or commodity chemicals. Unfortunately, photosynthetic organisms are not always ideal host for the production of desired chemicals and are frequently difficult to engineer. In order to bypass those hurdles, this work focused on introducing the machinery responsible for the light-energy conversion into a nonphotosynthetic host. The supplementation of a heterologous host with the energy captured via the light-energy conversion could alleviate some of the host's metabolic burden and allow for greater yields of desired compounds.

In order to achieve our goals, we set out to engineer functional expression of the bacterial reaction center from *R. sphaeroides* as well as the enzymes required for the production of bacteriochlorophyll into *E. coli*. For the first time we were able to demonstrate the expression of the reaction center complex as well as its primarily polar localization with *E. coli* cells. Furthermore, we characterized two previously poorly understood enzymes involved in the production bacteriochlorophyll, the 8-vinyl reductase (BciA) and the Mg protoporphyrin monomethylester cyclase (BchE). In the case of BciA, we showed that unexpectedly the BciA from *R. sphaeroides* was not functional when expressed in *E. coli,* unlike the BciA from *C. tepidum.* At the beginning of this work, BchE was the only enzyme involved in the biosynthesis of bacteriochlorophyll that has

not been heterologously expressed and had no published biochemical or biophysical data. Through our efforts, we were able to demonstrate that BchE contained an oxygen sensitive 4Fe-4S cluster able to interact with SAM, the predicted co-factor. Additionally, for the first time, we showed the interaction of BchE with several intermediates of the bacteriochlorophyll biosynthetic pathways.

Complementary to our efforts, we also produced a set of protein expression vectors for use in *R. sphaeroides*. *R. sphaeroides* is a photosynthetic organism which has been used extensively for the production of value added compounds and has the potential to be used for the production of membrane proteins. The novel vectors are BioBrick™ compatible and contain DsRed as a reporter protein driven by the photosynthetic *puf* promoter. We demonstrated that by selecting which section of the promoter was utilized in combination with various culture conditions, final reporter levels could be modulated. Reporter levels ranged from virtually undetectable to higher than what is present in *E. coli* when expression is driven from a constitutive *lac* promoter from the same vector backbone.

Table of Contents

Contents

List of Tables

List of Figures

Chapter 1. Introduction

This chapter is adopted from: Synthetic Biology: Tools and Applications. Chapter 16: Towards Engineering Light-Energy Conversion in Nonphotosynthetic Microorganisms. Ilya B. Tikh and Claudia Schmidt-Dannert.

Photosynthesis has been essential for the development of complex life on earth and is generally defined as the ability to convert light into chemical energy coupled to carbon fixation. Depending on their needs, organisms have developed different strategies for capturing and utilizing energy from the sun (1). The diversity of light capturing machinery ranges from the simple proteorhodopsin, to the incredibly complex multi-protein assemblies of photosystems I and II found in plants and cyanobacteria (1).

Photosynthetic organisms initially store produced energy in the form of a proton gradient across a membrane. However, the localization of that membrane varies between organisms (2). In most bacteria, such proton gradient is created across the inner membrane (2). In more complex photosynthetic systems, a special organelle, such as a thylakoid in cyanobacteria, houses all of the photosynthetic machinery and is used to generate a proton gradient (3, 4). After a proton gradient is created, its energy can be converted to chemical energy and stored in the form of ATP or NAD(P)H, which are subsequently utilized by the organism for growth (5, 6). Ultimately, some of the captured energy ends up as C-C bonds during CO_2 fixation.

Unlike truly photosynthetic organisms, phototrophic organisms generally use a simpler method for light capture and are not able to utilize CO_2 as their sole carbon source (7). Instead of CO_2 fixation, these organisms use the additional energy from light capture to help drive other metabolic pathways, reviewed in (1).

It is important to note that the core photosystem components responsible for light capture and conversion do not exist in a vacuum. Light-energy conversion in a photosynthetic organism requires a number of accessory proteins (1, 8). Complex pigments like chlorophylls and carotenoids are also required for proper function of bacterial photosynthetic reaction centers (RCs) and photosystems I and II (PSI, PSII) from plants and algae (9, 10). The enzymes responsible for the production of these pigments are encoded in large operons and their importance will be discussed later (11, 12).

Introduction to photosynthetic machinery as a way to power the cell.

Proteorhodopsin

The simplest and best characterized of the currently known mechanisms for light-energy conversion involves the rhodopsin protein family which is found in all kingdoms of life. Rhodopsins are trans-membrane proteins containing a single, light responsive retinal co-factor (Figure 1). Depending on the host, rhodopsins can function as light sensors, proton pumps, or ion pumps (reviewed in (13-15)).

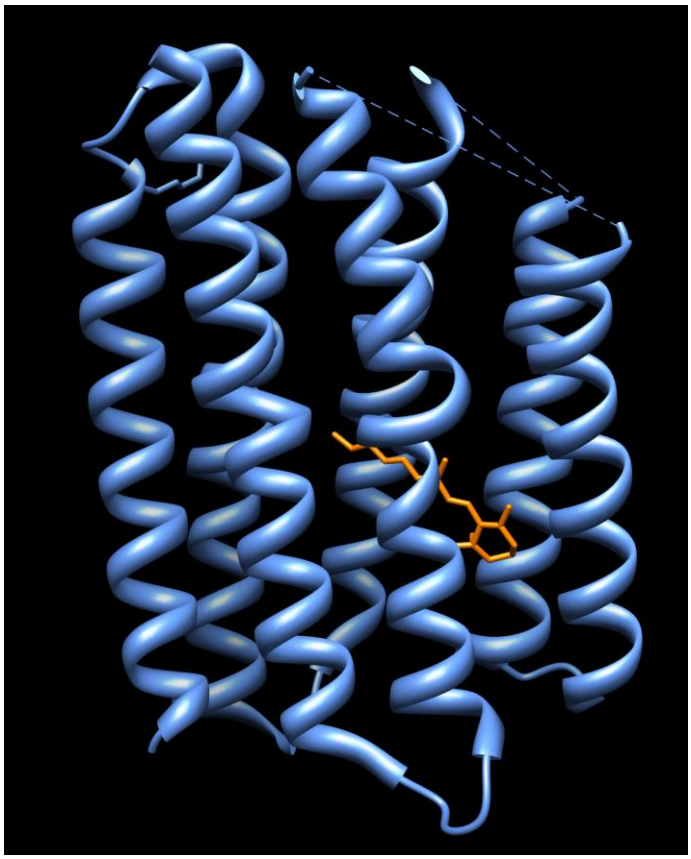

Figure 1.
3D structure of proteorhodopsin. Proteorhodopsin (PDB: 4JQ6) with the bound retinal cofactor shown in orange. (Ran *et al.* Acta Crystallograph, 2013)

Various members of the Rhodopsin family have been discovered and characterized. Bacteriorhodopsin from the archaeon *Halobacterium salinarium* was the first light-driven proton pump discovered in the early 1970's (16). To date, many archaeal bacteriorhodopsins have been identified from a wide range of environments using metagenomic sequencing techniques, with bacteriorhodopsin homologues being especially prevalent in the oceans, where resources are sparse, and any boost in energy generation can give an organism a significant competitive advantage. The first rhodopsin

homologue from a proteobacteria, proteorhodopsin, was identified is a metagenomic sample from the Sargasso Sea (17).

Recent work on light-driven proton pumping has centered around proteorhodopsins as opposed to bacteriorhodopsin, as the proteobacterial membrane proteins proved much easier to express and isolate from a wide range of heterologous hosts (17). Both bacteriorhodopsin and proteorhodopsin work by translocating a proton from the cytoplasm into the periplasm during the light induced isomerization of a retinol cofactor (14). Overall, this creates a very simple mechanism of establishing and maintaining a proton gradient.

Being such a simple system, proteorhodopsin is an obvious candidate to add into a recombinant host as a first step in engineering a heterologous light-energy conversion system. However, successful addition of proteorhodopsin to a new host requires not only proper folding and transmembrane localization of the protein, but also the availability of retinal as co-factor. Retinal can either be supplemented in the growth media or alternatively, a heterologous retinal biosynthetic pathway can be engineered for endogenous cofactor synthesis. For example, in *E. coli*, four additional gene products are needed to convert the isoprenoid precursor farnesyl diphosphate into β-carotene, which is then cleaved by a fifth enzyme at the central 15,15'-double-bond to yield retinal (18).

Since the discovery and characterization of proteorhodopsin, several groups have attempted to use it to supplement a heterologous host's energy pool (19, 20). There is evidence that expression of proteorhodopsin can aid in the regeneration of ATP in the host, especially under starvation or stress conditions. Studies have shown that starved *E.*

coli cells that express proteorhodopsin and are illuminated by light have more ATP than similar cells that are held in the dark (20). Furthermore, illumination of *E. coli* cells expressing proteorhodopsin can help negate the effects of respiration inhibition by azide (19).

In an attempt to take advantage of the increased proton pumping, proteorhodopsin was recently coexpressed with a hydrogenase in *E. coli* (21). In theory, the presence of proteorhodopsin coupled with a hydrogenase should offset some of the metabolic burden on the cell during hydrogen production and increase final H_2 yields by *E. coli* (21). As predicted, the addition of proteorhodopsin in that system did improve the final H_2 yields, though it is still not completely clear if the higher yield came from more protons being available for the hydrogenase or from a proteorhodopsin generated boost to the available energy levels inside the cell. If extra energy available via proteorhodopsin was the cause of increased H_2 production, other interesting uses can be envisioned such as the coupling of proteorhodopsin to carbon fixation in a non-photosynthetic host used in biofuels production. Fixing CO_2 would not only reduce feed stock requirements of the host, but it would also help create a more carbon neutral fuel product. But is proteorhodopsin able to generate enough energy to facilitate carbon fixation? A recent report shows that the presence of proteorhodopsin increases CO_2 fixation in the non-photosynthetic, marine bacterium *Polaribacter* sp. MED152, though the bacterium does require organic material in the medium for growth (22). It is still unclear if expression of proteorhodopsin in other non-photosynthetic hosts will generate sufficient energy to drive central metabolism, carbon fixation or other potentially useful metabolic reactions.

Several studies have shown that bacteria benefit the most from the presence of proteorhodopsin if they are under starvation, resource limiting conditions, or alternatively, if the proton gradient across the inner membrane has been disrupted (7, 23). There are several reasons why one could expect limited help from the addition of proteorhodopsin to an energy intensive metabolic process. One reason is the fairly low membrane potential generated by proteorhodopsin compared to the membrane potential maintained by *E. coli* under normal growth conditions (19). If proteorhodopsin could be engineered to generate a higher membrane potential, it would allow for an increased generation of ATP under all conditions, and not primarily when the host's respiration is otherwise compromised.

Introduction to bacterial reaction centers and their connection to the cell metabolism.

While proteorhodopsin is small and easy to heterologously express, its light-capturing efficiency is low compared to the reaction center. The greater efficiency of light-energy conversion by the bacterial reaction centers (RCs) compared to proteorhodopsin comes at a price of increased complexity. Bacterial RCs are multi-protein complexes containing a large number of pigments including bacteriochlophylls (Bchl) and carotenoids. Depending on the organism, either type I or type II RCs are present (Figure 2A). In the case of algae and cyanobacteria, both type I and II RCs make up the photosynthetic apparatus (1). Type II RCs, best studied in the purple non-sulphur bacteria, are cyclic, non-oxygen evolving photosystems. The high energy electrons generated by type II RCs are shuttled through the quinone pool, the cytochrome bc_1

complex and are eventually returned to the special Bchl pair of the RC via cytochrome c_2 (24). In contrast, type I RCs found in green-sulfur bacteria utilize ferrodoxin (Fd) as the final electron acceptor (3). Since the electrons do not return to the special Bchl pair of the RC, they need to be regenerated from a different source, frequently from oxidation of sulfur compounds. The majority of bacterial RCs are surrounded by light harvesting complexes containing additional pigments that increase the efficiency of light capture (Figure 2) (1).

Over the past several decades, the photosynthetic RCs several different *Rhodobacter* species have been used as models for the study of light capturing systems due to their relatively simple purification and stability after purification; thus making them most amenable for introduction into a recombinant host (25). Because of their stability, the core RCs, as well as the surrounding light harvesting complexes, have been thoroughly characterized. The RC in *R. sphaeroides* consists of three core proteins, referred to as the M, L and H subunits which are encoded by *pufM*, *pufL* and *puhH*, respectively. This core complex with its bound Bchl and carotenoid pigments, is surrounded by a light-harvesting (LH) superstructure formed by the LH1 and LH2 protein complexes that contain accessory Bchl pigments (Figure 2B,2C) (26).

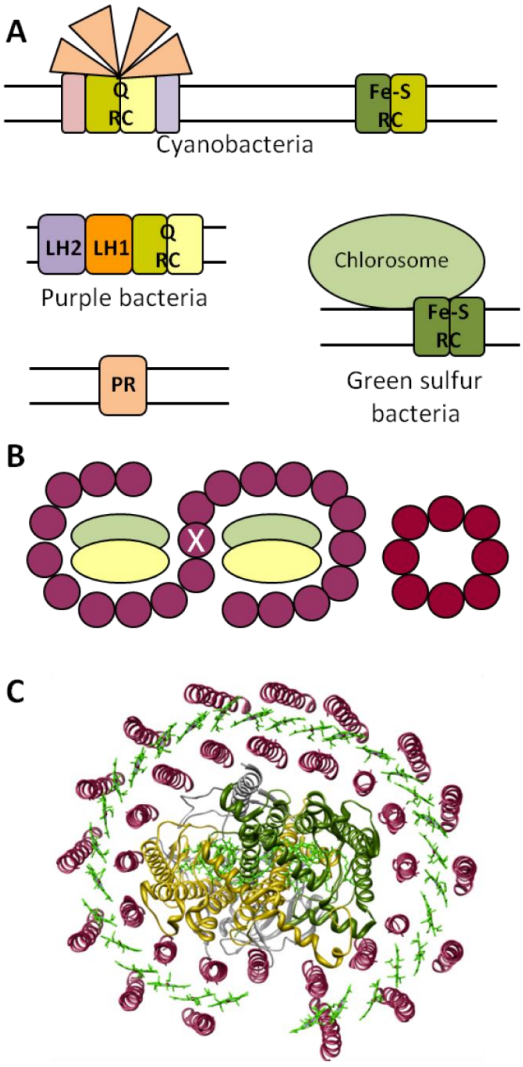

Figure 2.

Overview of composition of light harvesting machinery across different organisms. (A) Overview of the basic organization of the core reaction center (RC) complexes and the light harvesting subunits (LH). Colors of the core complexes indicate whether they are composed of homo-dimers or hetero-dimers. Q or Fe-S indicate quinones or iron-sulfur clusters (respectively) as the pathway for the transfer of excited electrons from the special pair to the final acceptor (modified from Bryant *et. al.*). **(B)** Schematic representing the organization of the RC-LH1-LH2 complex in *R. sphaeroides* as is present in the plane of the membrane. The core L, M and H subunits are (green and yellow) are surrounded by the LH1 complex (purple). The LH2 complex (red) is localized to the periphery of the RC-LH1 complex. **(C)** Top view of a crystal structure of the core RC-LH1 complex from *R. capsulatus* (PDB# 1PYH). The core L, M and H subunits are in located in the center surrounded by the αβ-peptides composing the LH1 (purple). Bacteriochlorophyll molecules are shown in green. The structure illustrates the complexity and precise assembly of the light-capturing machinery.

The function of the LH1 and LH2 complexes is to funnel captured light energy towards the core RC complex. After light strikes the reaction center, it excites a low energy electron from one of the two core Bchla molecules (referred to as the "special pair") located at the center of the type II RC complex. The high energy electron is transferred through the RC via two additional chlorophyll pigments and a quinone to a final quinone electron acceptor that exits the RC core upon reduction by two electrons. Thus, two excitation reactions are necessary before a quinone molecule is released from the RC complex (reviewed in (27)). The high energy e^- are cycled from the quinone back to the RC through the electron-proton transport chain and ultimately cytochrome c_2 (6). Proton translocation across the inner membrane is coupled to the decrease in the potential energy of the excited electrons during each step of the cyclic electron transfer. The established proton gradient can then be utilized by proton-gradient coupled enzymes such as ATP synthase to regenerate the metabolic needs of the cell (19). Alternatively, reduced quinones can be used directly by enzymes such as NADH:quinone oxidoreductase to regenerate NADH reducing equivalents needed for CO_2 fixation or other metabolically taxing reactions (Figure 3) (3).

In the non-cyclic, type I RC mentioned above, the high energy electrons are placed on a ferrodoxin molecule instead of a quinone. The ferrodoxin is subsequently released and is used as an energy source to convert $NAD(P)^+$ to NAD(P)H by the ferrodoxin-NAD reductase (Figure 3) (3). Similarly, the electrons in the special pair are regenerated from a c-type cytochrome; though this time the electrons originate from a different donor, frequently a sulfur compound (3).

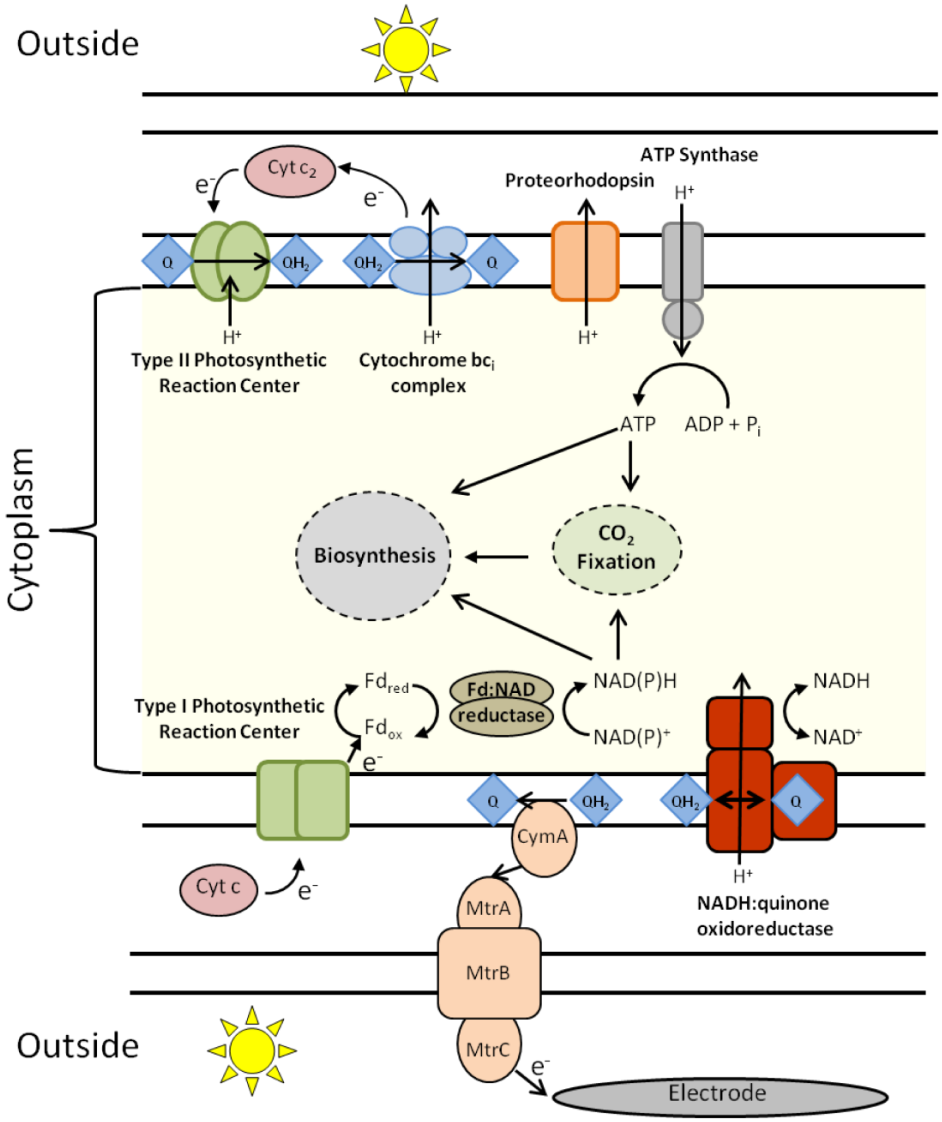

Figure 3.

Integration of light harvesting components into cell metabolism. Top portion of the figure demonstrates how a proton-translocating light-capturing system such as proteorhodopsin (PR) or a type II reaction center would integrate with the host's ATP Synthase, providing a boost to cell energy stores. Upon light excitation, PR is able to transfer a proton from the cytoplasm to the periplasm using the isomerization of a retinol cofactor. Type II RC does not translocate protons directly. Two protons bind to the reduced quinone on the cytoplasmic side of the membrane and are released on the periplasmic side during the oxidation of the quinone. The cytochrome bc_1 complex, responsible, which couples the oxidation of quinones to the reduction of cytochrome c2, also translocates protons. The bottom section illustrates how a non-cyclic type I RC can be used by the cell to regenerate reduced equivalents such as NADPH. After light excitation, electrons are transferred onto a ferrodoxin (Fd) molecule, which can be subsequently oxidized by Fd:NAD reductase in order to regenerate NADH. The transfer of electrons from the quinone pool through the *S. oneidensis* MR1 metal reduction pathway to an external acceptor, such as an electrode, is also shown in the bottom section of the figure.

The extra energy boost generated by either the type I or type II RC would be ideally coupled to energy expensive metabolic reactions such as the synthesis of long chain fatty acids which are readily transformed into biofuels. Fatty acid synthesis utilizes malonyl-CoA molecule as a basic building block and requires hydrolysis of ATP to create C-C bonds (reviewed in (28, 29)). If a cell is engineered for overproduction of fatty acids, any potential boost to its total ATP pool should prove beneficial to the final product yields. Not only are ATP molecules needed during the production of long chain fatty acids, but a number of NADPH molecules are also used up in the process of making a saturated fatty acid. As discussed previously, NADPH could be regenerated by the use of reduced quinones or from ferredoxin molecules generated during electron transfer from the reaction center (Figure 3) (3, 24).

Introduction to the pigments involved in photosynthesis

As mentioned previously, (bacterio)chlorophylls and carotenoids are essential for proper function of the light harvesting apparatus in both plants and bacteria. Their interactions with the photosynthetic machinery as well as their biosynthesis have been extensively studied. The presence of both (bacterio)chlorophylls and carotenoids is as essential for photosynthesis as the reaction center itself.

Carotenoids

Carotenoid pigments serve as accessory pigments in the core RC complex as well as in light-harvesting complexes surrounding it. Their three primary functions, discussed below, include photoprotection of the reaction center, harvesting and transferring light energy to the chlorophyll molecules and, in certain cases, contributing to the stability of

the photosynthetic components (reviewed in (30)). The structure of carotenoids present in the photosynthetic machinery is widely varied between the organisms; some of the well described carotenoids are shown in Figure 4. In majority of instances, carotenoids biosynthesis begins with head-to-head condensation of two geranylgeranyl pyrophosphate molecules by phytoene synthase (CrtB) to produce phytoene (reviewed in (31)). Phytoene is subsequently desaturated and otherwise modified by enzymes contained in the crt operon to produce the final carotenoids (31). The modifications of the carotenoids backbone frequently differ between organisms, for example, cyanobacteria will frequently cyclize one or both ends of the lycopene molecule, where as carotenoids produced by purple bacteria are rarely cyclized.

Photoprotection

Photoprotection is generally regarded as the primary function of carotenoids and *Rhodobacter* strains not making carotenoids were found to be significantly more susceptible to damage by high illumination or oxygen concentrations (32). It was later discovered that in the absence of carotenoids, triplet-excited bacteriochlorophylls would react with O_2 to produce singlet O_2, a powerful oxidizing agent which would quickly kill the cells. Carotenoids are able to quench the triplet bacteriochlorophyll state significantly faster than molecular oxygen, thus reducing oxidative stress on the cells (30, 33).

Photon Harvesting

In addition to photoprotection, some carotenoids are able to absorb light in the 450-550nm range and subsequently transfer some of that energy onto the nearby chlorophyll molecules (reviewed in (30)). The ability to capture light in that range is

especially useful for anaerobic phototrophs as they frequently have access to a limited amount of light since the aerobic phototrophs reside closer to the surface and absorb light in the 600-700nm range. Energy transfer from carotenoids to (bacterio)chlorophylls is inefficient with only 30-60% of the captured energy making it to the acceptor (bacterio)chlorophyll (34). While this may represent a small boost to the very efficient energy transfer between the light harvesting complexes and the reaction centers, it could prove essential because of the competition for resources.

Structural component

Since carotenoids are such an integral part of photosynthetic machinery, it is understandable that they play a role in its assembly. Structural studies have revealed binding sites for carotenoids in both the reaction centers and the peripheral light harvesting complexes (35, 36). While the carotenoid binding pocket in the reaction center can be filled by a lipid or detergent without affecting the overall structure of the complex, carotenoids were found to be essential for the stability of the light-harvesting LH2 complex in *R. sphaeroides* (35).

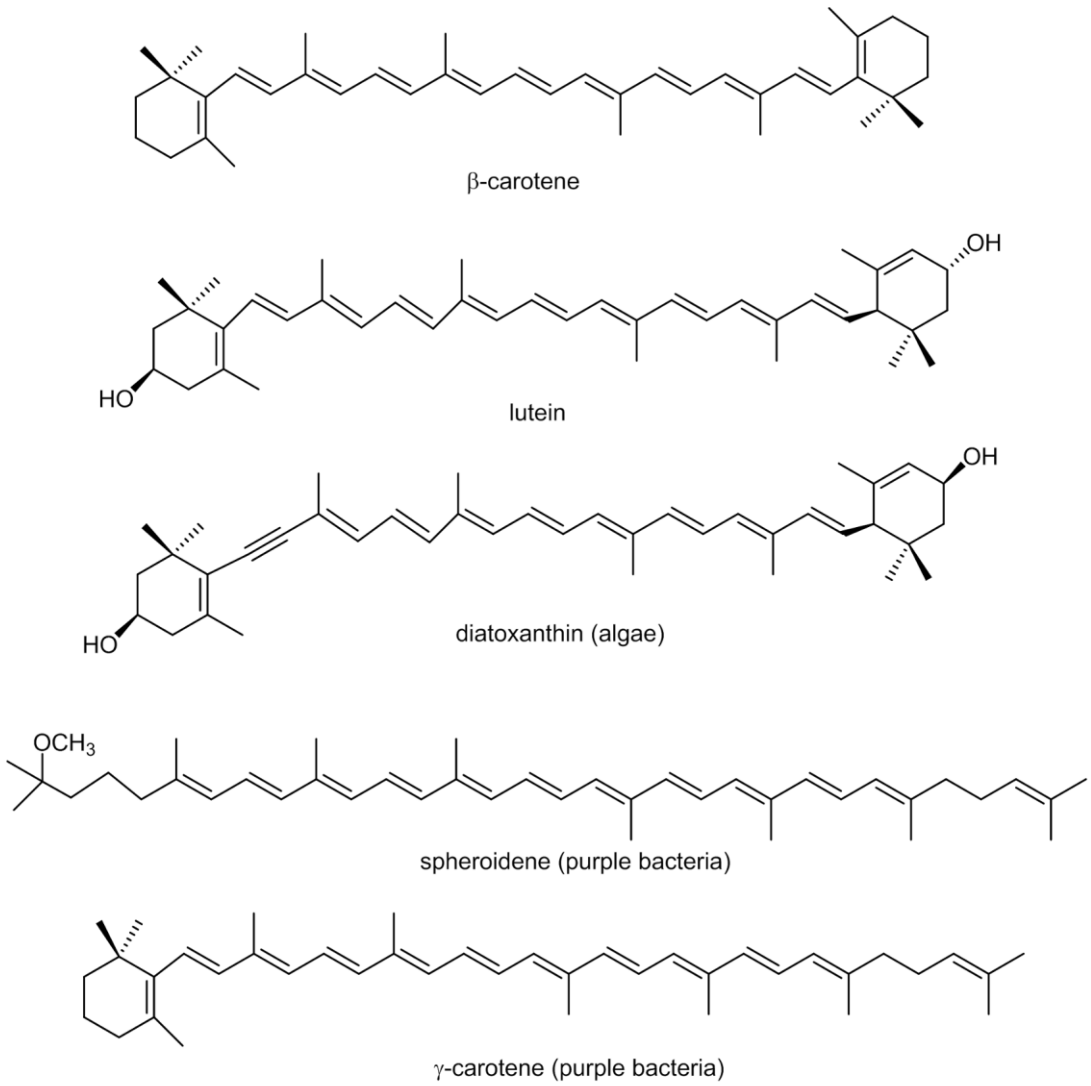

Figure 4.
Overview of several common carotenoids found in plants, algae and photosynthetic bacteria.

Bacteriochlorophylls.

Biosynthesis of (bacterio)chlorophyll is known to require over 20 different enzymatic steps starting from succinyl-CoA and glycine as the precursons of the heme biosynthetic pathway (reviewed in (37)). The first committed step in Bchla biosynthesis is the chelation of magnesium into protoporphyrin IX by the BchHID enzyme complex (Figure 5) (38). Subsequently, BchM methylates the carboxyl group of the 13-propyonate on magnesium protoporphyrin IX in a SAM dependent fashion (38, 39). Next, a cyclization step is performed to create the characteristic fifth ring common to all (bacterio)chlorophylls (40). The D-pyrrole ring of the newly created 3,8-divinyl protochlorophyllide is reduced by either a light dependent (LPOR) or light independent (DPOR) enzyme complex (reviewed in (41)). At this point, depending on the final type of (bacterio)chlorophyll to be produced, some organisms utilize a divinyl-reductase to convert the 8-vinyl chlorophyllide a to 8-ethyl chlorophyllide a (42-44). The produced chlorophyllide a then acts as the last common precursor for the production of various types of (bacterio)chlorophylls (Figure 5) (reviewed in (41)).

While the enzymes involved in the biosynthesis of Bchl have all been identified, two of the enzymatic steps in the process have yet to be biochemically characterized. The first of these is the formation of the E-ring on the porphyrin backbone. In fact, two potential enzymes, BchE and AcsF, have been identified as candidates to catalyze the cyclization reaction of MgPIX ME (40, 45, 46). Little is known about either of the two enzymes, except that strains containing deletions of those genes accumulate Mg-protoporphyrin IX monomethylester (40). BchE is primarily expressed under anaerobic

conditions and contains a well conserved cobalamin binding domain in addition to the CXXXCXXC motif associated with 4Fe-4S clusters. Even less is known about AcsF it contains a consensus motif (D/E)EXXH, frequently found in metal containing monooxygenases (40, 47). Cell fraction studies performed with barley and cucumber suggest that a membrane associated component is needed for correct function of plant AcsF (46, 48).

The second poorly characterized step in the (bacterio)chlorophyll biosynthetic pathway comes at the point of the reduction of the C-8 vinyl group. Genetic studies initially implicated BchJ as the enzyme responsible for the reduction of 3,8-divinyl-protochlorophyllide, however, recent evidence suggest at least two other enzymes perform that task (41, 44, 45, 49, 50). The newly identified 8-vinyl reductases have been named BciA and BciB and both are well conserved across most kingdoms of life, however, they appear to be evolutionary unrelated to each other. Both of the enzymes were initially identified as 3,8-divinyl-protochlorophyllide reductases, however, *in vivo* experiments have produced mixed results about the preferred substrate for these enzymes as they are able to reduce substrates from MgPIX to divinyl-chlorophyllide a. Purification and proper biochemical characterization of BciA and BciB should be able to shed some light on true substrate specificity and hopefully their placement in the (bacterio)chlorophyll biosynthetic pathway.

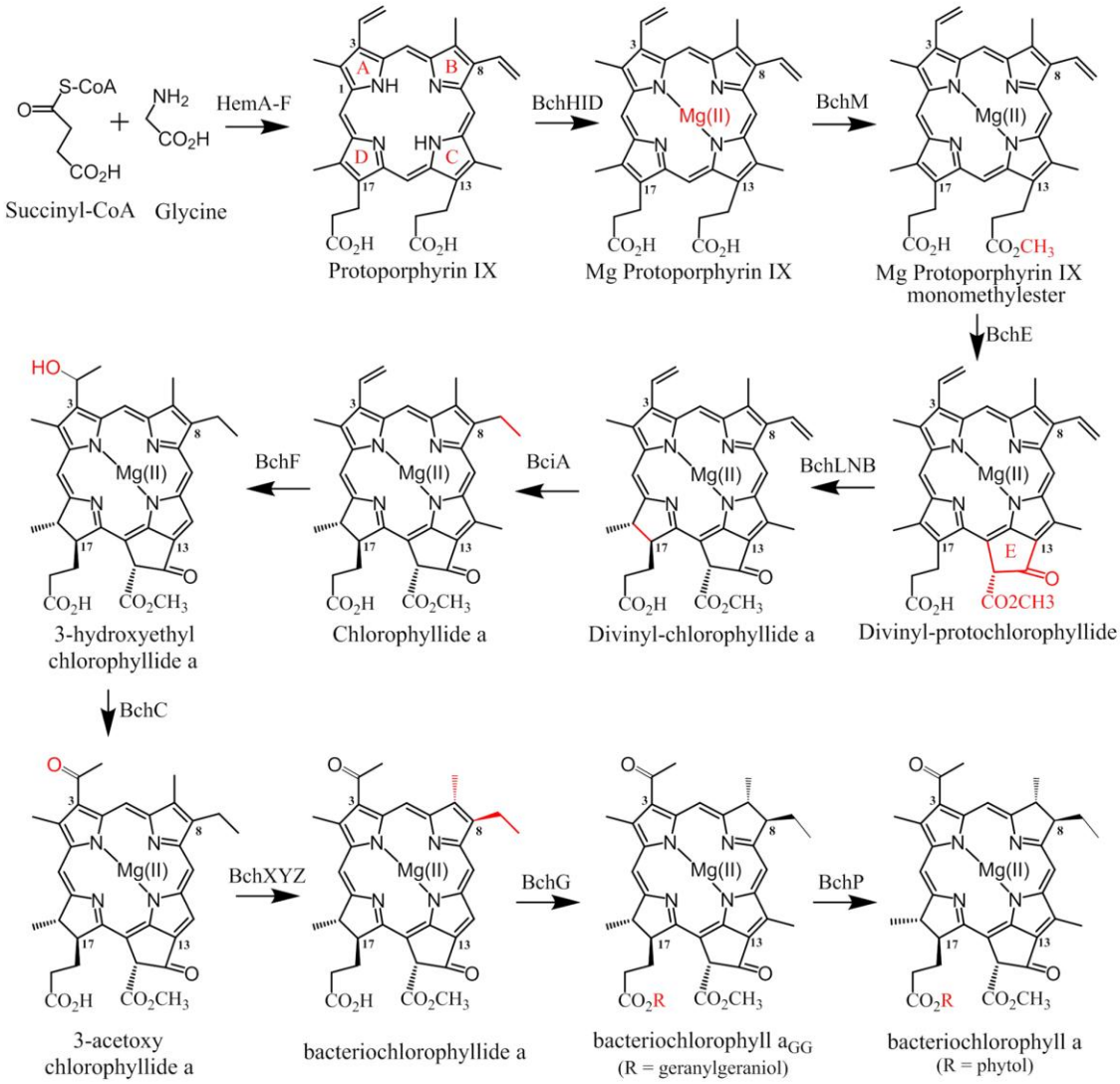

Figure 5.

Bacteriochlorophyll biosynthetic pathway. Pathway shown from the beginning of the heme biosynthetic pathway utilizing succinyl-CoA and glycine as starting points and continuing through heme biosynthesis (HemA-F). Protoporphyrin IX serves as the last common intermediate between the heme and bacteriochlorophyll biosynthetic pathways. Changes to the porphyrin backbone during each step are highlighted in red.

Chapter 2. BioBrick™ compatible vector system for protein expression in *Rhodobacter sphaeroides.*

This chapter is adopted from: BioBrick™ compatible vector system for protein expression in *Rhodobacter sphaeroides*. Ilya B. Tikh, Mark Held and Claudia Schmidt-Dannert. Submitted for publication in Applied Microbiology and Biotechnology.

Chapter 2 Overview

We report here the development of a modular, plasmid-based protein expression system utilizing elements of the native *Rhodobacter puf* promoter in a BioBrick™ based vector system with *DsRed* encoding a red fluorescent reporter protein. A suite of truncations of the *puf* promoter were made to assess the influence of different portions of this promoter on expression of heterologous proteins. The 3' end of *puf* was found to be particularly important for increasing expression, with transformants accumulating significant quantities of DsRed under both aerobic and anaerobic growth conditions. Expression levels of this reporter protein in *R. sphaeroides* were comparable to those achieved in *E. coli* using the strong, constitutive P_{lac} promoter, thus demonstrating the robustness of the engineered system. Furthermore, we demonstrated the ability to tune the designed expression system by modulating cellular DsRed levels based upon the promoter segment utilized and oxygenation conditions. Lastly, we show that the new expression system is able to drive expression of a membrane protein, proteorhodopsin and that membrane purifications from *R. sphaeroides* yielded significant quantities of proteorhodopsin. This toolset lays the groundwork for the engineering of multi-step pathways, including recalcitrant membrane proteins, in *R. sphaeroides*.

INTRODUCTION

Rhodobacter sphaeroides is a purple, non-sulfur, photosynthetic bacterium, which has been a model organism for the study of photosynthesis for over three decades (1). In addition to the photoautotrophic growth, *R. sphaeroides* has a highly diverse metabolism; enabling it to grow under a wide range of both aerobic and anaerobic conditions. This metabolic versatility makes *R. sphaeroides* an attractive candidate for industrial applications (2-4). For example, carbon dioxide is readily utilized by this bacterium and thus, it is easy to envision utilizing *R. sphaeroides* for the production of carbon neutral bioproducts (5, 6). Furthermore, recent studies have shown that *R. sphaeroides* can be used to produce hydrogen gas from a range of carbon sources, including organic waste (2, 7-10).

During photoautotrophic growth, *R. sphaeroides* has the capability of forming extensive intracytoplasmic membranes (ICMs) (11, 12) The extra membrane space produced by the ICMs can be harnessed for the accumulation of value-added hydrophobic compounds, such as carotenoids or quinones (3, 13), or the production of membrane proteins (14, 15). Clearly, *R. sphaeroides* is a versatile host suitable for a wide range of applications including protein expression, biotransformation and bioproduction.

Unfortunately, advanced engineering of this bacterium is hampered by a lack of sophisticated genetic tools, which are vital for any biotechnological application. Most reports attempting genetic modifications have focused primarily on chromosomal manipulation of *R. sphaeroides* for research related to the photosynthetic machinery. Although genomic insertions and deletions can address some very specific questions, the

process is often slow and tedious (16, 17). Additionally, unwanted and potential deleterious polar effects on downstream genes are a common side-effect of this strategy.

Transformation of *R. sphaeroides* to address the topics mentioned above has been achieved through the development of several plasmid systems, the vast majority of which are designed for genetic complementation, not high level protein expression, thus limiting their potential for biotechnological applications (14, 18, 19). Not surprisingly, these plasmids used in *Rhodobacter* tend to be large, without optimized cloning sites or well characterized promoters. All of these limitations make sophisticated genetic engineering in *R. sphaeroides* cumbersome.

In this study we sought to expand the genetic toolkit available for this versatile purple bacterium by creating a set of convenient vectors that drive protein expression from the so-called *puf* promoter (referred to P_{puf} hereafter). P_{puf} resides upstream of the *puf*B gene, at the 5' of the *puf*-operon, which encodes the structural genes of the light-harvesting LH-1 complex and the photosynthetic reaction center (RC) (20). Our engineered vectors were built off the pBBRBB-BioBrick™ plasmid previously developed in our laboratory (available from www.addgene.org) and allow stacking of multiple gene expression cassettes as BioBricks™ (21-23). Promoter function in *R. sphaeroides* was verified by monitoring the expression of the red fluorescent reporter protein DsRed Express2 (hereafter referred to as DsRed)(24). Finally, utility of the created modular expression system for heterologous production of membrane proteins in *Rhodobacter* was demonstrated by overexpression and isolation of the model membrane protein proteorhodopsin.

MATERIALS AND METHODS

Enzymes and chemicals

All enzymes were purchased from New England Biolabs (NEB) and used in accordance with supplied instructions. Chemicals used in the study were purchased from Sigma-Aldrich, unless otherwise noted.

Cell growth conditions

Strains and plasmids used in this study are listed in Supplementary Table 1. *E. coli* cultures used were grown at 37^{o}C in Luria-Bertani (LB) medium supplemented with 30 μg/ml kanamycin (MP Biomedical) for plasmid selection as needed. *E. coli* JM109 (DE3) (Promega) were utilized for all of the molecular biology work, *E. coli* BL21 (DE3) were utilized for the expression of DsRed. *Rhodobacter* strains (wild-type strain 2.4.1 and deletion strain ΔRCLH) were grown at 30^{o}C in 4 ml of either LB or Sistrom's minimal medium (25) in 16x100 mm culture tubes supplemented with kanamycin as needed. High culture oxygenation was achieved by using 1 ml of growth medium in 25x150 mm culture tubes with shaking at 250 rpm. Plasmid conjugation into *Rhodobacter* was performed as described by Saltikov and Newman (26) using *E. coli* WM3064 for plasmid conjugation. Anaerobic cultures were made by bubbling N_2 gas through 10 ml of medium in 18x150 mm tubes for 10 min prior to sealing tubes with butyl stoppers and autoclaving. Kanamycin, dimethyl sulfoxide (DMSO) and culture inoculums were added via gas tight syringe post autoclaving. During anaerobic growth in Sistrom's medium, succinate (30 mM) and DMSO (50 mM) were added to the culture to serve as an electron donor and acceptor, respectively.

Plasmid constructions

P_{puf} promoter fragments were amplified from *R. sphaeroides* genomic DNA with the primers indicated in Table 1. *R. spharoides* genomic DNA was isolated from *R. sphaeroides* 2.4.1 (ATCC number 55304) using the Wizard genomic DNA isolation kit (Promega) according to manufacturer's instructions. For cloning, the pBBRBB-DsRed vector was created by amplifying DsRed Express 2 (24)(hereafter referred to as DsRed) from pUCBB-DsRed (23) using primers P1 and P2 containing *Bgl*II and *Not*I restriction sites. The amplified product as well as pBBRBB-GFP were both digested with *Bgl*II and *Not*I. Subsequently, DsRed was ligated into the pBBRBB backbone creating pBBRBB-DsRed. The P_{puf} promoter segments were amplified with primers P2, P3 ($P_{puf1\text{-}1200}$), P4, P5 ($P_{puf1\text{-}842}$) and P6, P7 ($P_{puf843\text{-}1200}$) and digested with *Xba*I and *Bgl*II.The digested products were ligated into pBBRBB-DsRed cut with the same enzymes, ligated using T4 DNA ligase (NEB) and transformed into *E. coli* JM109(DE3) to generate pBBRBB-$Ppuf_{1\text{-}1200}$, pBBRBB-$Ppuf_{1\text{-}842}$ and pBBRBB-$Ppuf_{843\text{-}1200}$.

Site directed mutagenesis to remove the *Eco*RI and *Pst*I sites from pBBRBB-$Ppuf_{1\text{-}1200}$ and pBBRBB-$Ppuf_{1\text{-}842}$ was performed using the QuickChange™ (Agilent) protocol with primers P9, P10 and P11, P12, respectively (Table 1). The sequence of all constructs was verified by DNA sequencing. Plasmids will be deposited upon publication (language will be changed to "are deposited") to the not-profit plasmid repository Addgene (www.addgene.org) where they will be available to the research community.

Table 1. Primers described in Chapter 2

P1	5'-atctaga-ATCGTCATGTACCGCGAATTCG-3'
P2	5'-agt-agatct-GCTATCCTCCGGATCGTAAGAC-3'
P3	5'-atctaga-TTCGGCGAGAGGAAGGGAGAG-3'
P4	5'-agt-agatct-TGGTTCTCTCCCTTCCTCTC-3'
P5 (Proteorhodopsin F)	5'-at-agatct-ATGAAATTATTACTGATATTAGGTAGTG-3'
P6 (Proteorhodopsin R)	5'-at-gcggccgc-ttaatggtgatgatggtgatg-AGCATTAGAAGATTCTTTAACAGC-3'
P7 EcoRI removal F	5'-CGTCATGTACCGCGAAcTCGGCCGCGGGCTGGCCG-3'
P8 EcoRI removal R	5'-CGGCCAGCCCGCGGCCGAgTTCGCGGTACATGACG-3'

Flow cytometry

Flow cytometry analysis was performed using a BD FACScalibur flow cytometer using a 560 nm excitation and 661 nm emission filters (FL4) at a low flow rate. Samples (from three independent cultures for each vector analyzed) were withdrawn from the cultures at 24, 48 and 72 hours post inoculation and were diluted 1000 fold into phosphate buffered saline (PBS) prior to analysis. Anaerobic cultures were allowed to sit at 4^{o}C in PBS for one hour prior to analysis in order for the fluorescent protein to oxygenate and fully mature. A total of 50,000 events were collected per sample at a rate of approximately 1,000 events per second. Flow cytometry data was analyzed using FlowJo version 7.5 (www.flowjo.com).

For *E. coli* cultures, three individual colonies from a plate containing a fresh transformation of BL21 (DE3) cells with pBBRBB-P_{lac}-DsRed were inoculated into LB and grown at 37°C to provide data in triplicate. For *R. sphaeroides* cultures, three individual colonies per plasmid were picked from a fresh conjugation plate were inoculated into 4 ml of LB containing 30 µg/ml of kanamycin. Subsequently, the cultures were allowed to grow for 72 hours to reach stationary phase. From stationary cultures, 100 µl was used as inoculum into 10 ml for anaerobic cultures and 50 µl was inoculated

into 4 ml for aerobic cultures, were used as inoculums for the flowcytometry experiments. In the case of high aeration cultures, 20 µl of stationary phase cells was used to inoculate 1 ml of culture.

Purification of proteorhodopsin

Proteorhodopsin (PR) (Genbank accession number AAG10475) was amplified from pBBR1MCS-2-PR (27) using P7 and P8, digested with *Bgl*II and *Not*I and ligated into pBBRBB-$Ppuf_{843\text{-}1200}$ digested with the same restriction enzymes to create pBBRBBB-$Ppuf_{843\text{-}1200}$-PR. A C-terminal 6x-His tag was introduced with the reverse primer P8. The resulting construct, pBBRBB-$Ppuf_{843\text{-}1200}$-PR was transferred via conjugation into both wild-type strain *R. sphaeroides* 2.4.1 and reaction center deletion strain *R. sphaeroides* ΔRCLH (16). For proteorhodopsin purification, a 48 hour 4 ml culture from a fresh conjugation plate was transferred into 1 L of LB in a 2 L flask supplemented with 10 µM all-trans retinal for reconstitution of the apo-PR protein (27). After 60 hours of growth at 30°C with shaking at 250 rpm, cells were harvested by centrifugation and resuspended in 25 ml of lysis buffer (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM imidazole) and lysed using sonication. After sonication, powedered n-Octyl-β-D-glucopyranoside (OG) (Gold Biotechnology) was added to the lysate to final concentration of 2% w/v and allowed to incubate for 1 hour at 4^{o}C with gentle stirring in order to solubilize membrane proteins. Soluble proteins were separated from the cell debris by centrifugation at 15,000 g for 45 min at 4°C, and the supernatant was further clarified by passing through a 0.45 µm filter device (EMD Millipore). The clarified, soluble fraction, approximately 20 ml,

was loaded onto a 5 ml Fastflow Ni+ column (GE Life Sciences) equilibrated with buffer A (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM imidazole, 0.2% N,N-Dimethyldodecylamine N-Oxide (LDAO)) and washed with two column volumes of buffer A. The protein was eluted using a gradient of 0-100% buffer B (50 mM Tris pH 8.0, 250 mM NaCl, 250 mM imidazole, 0.2% LDAO) over 100 ml. PR containing fractions (displaying the characteristic holo-PR absorption spectra and maximum at 520 nm) were pooled and concentrated. Retinal-bound holo-PR concentration was determined using a molecular extinction coefficient of 50000 $M^{-1}cm^{-1}$ at 520 nm. Protein concentration was determined using the Bradford assay (Pierce, Rockford, IL).

RESULTS

Creation of a BioBrick™ expression system for *R. sphaeroides*

To create a convenient, modular expression vector for *Rhodobacter* that conforms to the BioBrick™ standard used by synthetic biologists, we chose to start with the broad-host range pBBRBB-BioBrick™ vector previously developed in our laboratory (23). This plasmid contains a broad-host range origin of replication and can be maintained in a diverse range of bacteria using kanamycin for selection (28). In addition, it features mobility or *mob* genes for rapid plasmid transfer into multiple hosts via conjugation using *E. coli* strains such as WM3064 (26).

From the limited number of promoters (14, 16, 18, 19) available for heterologous gene expression in *Rhodobacter*, we chose the *puf* promoter region (20) to design a strong promoter for this genus. As described previously, the P_{puf} promoter in *R.*

sphaeroides drives the expression of the *pufBALMX*, which compose the light harvesting (LH1) complex, as well as the reaction center complex (Figure 1A). The *puf* promoter region is frequently defined as the 1200 bases upstream of *puf*B, and contains two open reading frames (ORFs) coding for small peptides PufQ and PufK, thought to regulate translation of the LH1 complex and the reaction center (14, 18, 20). Because little is known about which segments of the 1200 bp *puf* promoter region are required to initiate and/or control transcription in *R. sphaeroides*, we amplified three segments, $P_{puf1\text{-}1200}$, $P_{puf1\text{-}842}$ and $P_{puf843\text{-}1200}$ (Figure 6A). The $P_{puf1\text{-}1200}$ segment represents the full-length *puf* promoter region, containing 1200 base pairs upstream of the start codon of *puf*B, as well as the *puf*Q and *puf*K ORFs (Figure 6B). Conversely, $P_{puf1\text{-}842}$ contains the region upstream of *puf*Q and does not contain any portion of *puf*Q or *puf*K. Lastly, $P_{puf843\text{-}1200}$ contains only the *puf*QK ORFs. The constitutive P_{lac} promoter in pBBRBB-BioBrick™ was then replaced with these different sections yielding three expression vectors: pBBRBB-$Ppuf_{1\text{-}1200}$, pBBRBB-$Ppuf_{1\text{-}842}$ and pBBRBB-$Ppuf_{843\text{-}1200}$. Unfortunately, promoter sections $P_{puf1\text{-}1200}$ and $P_{puf1\text{-}842}$ contained *Eco*RI and *Pst*I restriction sites, which were removed by mutagenesis in order for the vectors to be compatible with the BioBrick™ standard. Removal of the sites did not affect promoter function as tested below (data not shown).

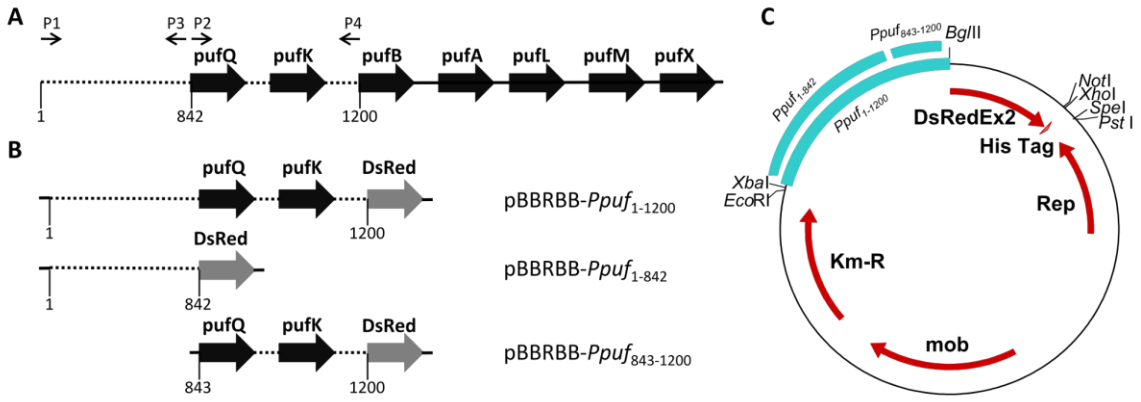

Figure 6.

Detailed map of the *R. sphaeroides puf* genomic region and BioBrick™ vectors constructed. A The full-length *puf* promoter (1200 bp in length, indicated by a dashed line) encompasses both the *puf*Q and *puf*K genes. P1 through P4 indicate the location and direction of primers used for amplifying the full-length P_{puf} promoter, the 1-842 section and the 843-1200 section. Downstream ORFs part of the puf operon including *puf*B, A, L, M, and X are also displayed. **B** Different segments of the P_{puf} promoter were transcriptionally fused to *DsRed* to assess promoter function in *R. sphaeroides* The expression of DsRed was compared between the full-length *puf* promoter, a 5' section containing base pairs 1-842, which lacks *puf*Q and *puf*K, and lastly, a 3' segment (base pairs 843-1200), which contains *puf*Q and *puf*K, but lacks the first 842 bps of the native promoter. **C** Combined plasmid map of the created vectors displaying the different promoter segments that drive expression of the *DsRed* reporter gene, the broad-host-range origin of replication (Rep), the mobilization region (Mob) and kanamycin resistance conferring gene (Kan-R). An in-frame, C-terminal 6X HIS-tag is available by cloning into the *Xho*I site.

Characterization of P_{puf} promoter function in *R. sphaeroides*

To assess the function of the three P_{puf} constructs *in vivo*, the red fluorescent reporter gene DsRed was transcriptionally fused to the three promoter regions as shown in Figure 1B. The expression level of DsRed in cultures was analyzed by flow cytometry under various growth conditions for both wild type *R. sphaeroides* and the ΔRCLH deletion strain. The ΔRCLH deletion removes the light-harvesting LH1 and LH2 complexes and the three genes encoding for the subunits photosynthetic reaction center (RC) (16). Deletion and removal of the photosynthetic machinery from *Rhodobacter's* intracytoplasmic membrane space is thought to free up metabolic resources for

heterologous protein expression in general and importantly, space for membrane protein expression.

Because the *puf* promoter is known to be repressed by oxygen (29), we first grew wild-type *R. sphaeroides* in minimal medium under anaerobic conditions. Our *puf* promoter constructs expressed noticeable levels of DsRed, well above that of the untransformed control (Figure 7A). Under these and indeed all conditions, cells harboring the pBBRBB-$Ppuf_{1\text{-}1200}$ and pBBRBB-$Ppuf_{843\text{-}1200}$ constructs (light blue and green traces, respectively) produced similar amounts of DsRed, while cells containing the pBBRBB-$Ppuf_{1\text{-}842}$ construct (lacking *puf*Q and *puf*K; orange trace) showed significantly lower DsRed levels. These results indicate that the level of expression from the *puf* promoter can be modulated, with the 3' region of P_{puf} being required for higher expression levels.

Next, we analyzed DsRed expression under the control of the three P_{puf} promoter constructs during aerobic growth conditions in minimal media (Figure 7C) and compared cellular DsRed levels to the anaerobic expression data. Surprisingly, similar levels of DsRed expression were noted between the same constructs under both conditions (compare panels 7C and 7A). This result was then recapitulated in rich LB media, as well (Figure 7E). These results were surprising considering the previously mentioned fact that the *puf* promoter is known to be repressed by oxygen. To address this inconsistency, we grew wild-type *R. sphaeroides* in larger volume culture tubes, containing very little medium as a means of increasing aeration and therefore oxygen concentration of these cultures. The increased aeration led to a significant repression of *puf* promoter activity

across all constructs, thus providing continuity between our work and previous reports on the subject (15, 29). Interestingly, the observed repression of DsRed expression was more pronounced in the two constructs ($P_{puf1\text{-}1200}$, $P_{puf1\text{-}842}$) that retain the region (1-842) upstream of *puf*Q.

We also analyzed the efficiency of the three P_{puf} expression constructs in the ΔRCLH deletion strain and found that all of the same trends observed with the wild-type strain were reiterated (Figure 7B, D, F, H). The only notable difference between the two strains was that the level of DsRed expression was consistently higher for the deletion strain matching our earlier hypothesis (compare Figure 7 left column to right column).

Finally, to get a sense of how *Rhodobacter* compares to *E. coli* as a host for heterologous protein expression using the P_{puf} promoter constructs, we transformed *E. coli* BL21 (DE3) with pBBRBB-P_{lac}-DsRed for expression of DsRed from the strong, constitutive P_{lac} promoter (Figure 2E, F, dark blue trace). Maximum levels of DsRed in *E. coli* were comparable to DsRed levels produced by *R. sphaeroides* 2.4.1 and actually lower than the DsRed levels in the *R. sphaeroides* ΔRCLH strains harboring the pBBRBB-$Ppuf_{1\text{-}1200}$ and pBBRBB-$Ppuf_{843\text{-}1200}$ constructs (Figure 7E, F). These results demonstrate the strength of our recombinant P_{puf} promoters.

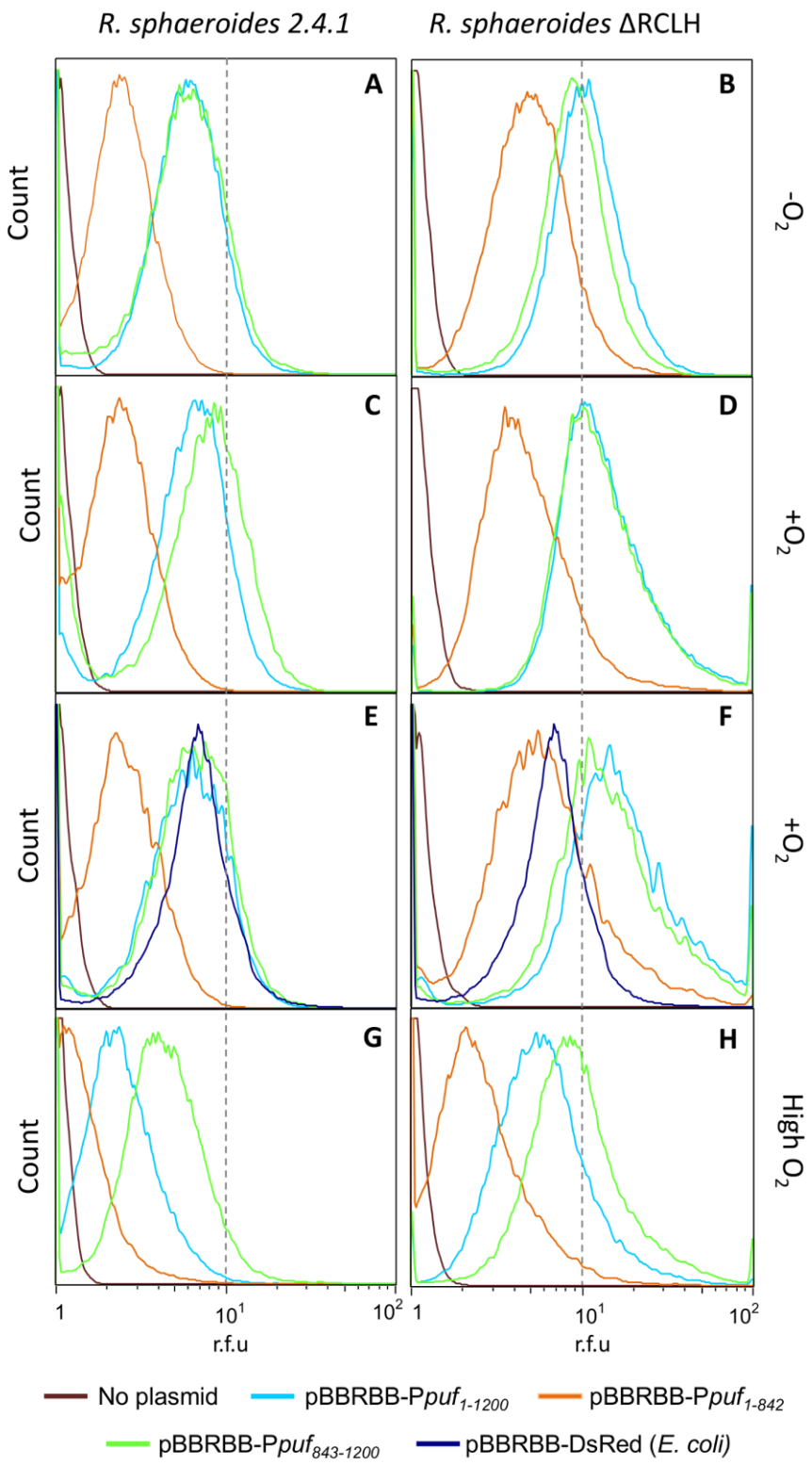
R. sphaeroides 2.4.1
R. sphaeroides ΔRCLH
A
B
C
D
E
F
G
H
Count
-O2
+O2
+O2
High O2
1
10^1
10^2
r.f.u
No plasmid
pBBRBB-Ppuf1-1200
pBBRBB-Ppuf1-842
pBBRBB-Ppuf843-1200
pBBRBB-DsRed (E. coli)

Figure 7.

Flow cytometry analysis of DsRed expression from engineered P_{puf} constructs. Wild-type *R. sphaeroides* 2.4.1 (left column) or ΔRCLH deletion strain cells (right column), transformed with the pBBRBB-P_{puf}-DsRed expression vectors were analyzed by flow cytometry after 48 hrs of growth at 30°C with varying growth media and oxygen levels. Transformants depicted are as follows: burgundy trace - untransformed control cells; light blue trace – pBBRBB-$Ppuf_{1\text{-}1200}$; orange trace – pBBRBB-$Ppuf_{1\text{-}842}$; green trace – pBBRBB-$Ppuf_{843\text{-}1200}$. **A** and **B** were grown anaerobically in Sistrom's minimal media. **C** and **D** were grown aerobically in minimal media. **E** and **F** were grown aerobically in LB medium. **G** and **H** were grown under high aeration conditions (see methods) in LB medium. For comparison of promoter strength, *E. coli* BL21 (DE3) cells were transformed with the expression vector pBBRBB-P_{lac}-DsRed, containing the strong, constitutive P_{lac} promoter and grown under aerobic conditions in LB medium (dark blue trace).

Application of P_{puf} BioBrick™ vectors for membrane protein expression

After confirming that our P_{puf} promoter segments were active in our engineered vector system, we sought to analyze how well *R. sphaeroides* would express recombinant membrane proteins. To that end, a model membrane protein, proteorhodopsin was cloned with a C-terminal His-tag into the pBBRBB-$Ppuf_{843\text{-}1200}$ vector. This vector was chosen because of its' high expression level compared to the pBBRBB-$Ppuf_{1\text{-}842}$ vector (Figure 7). Following 60 hours of aerobic growth in rich medium, proteorhodopsin was solubilized and purified via metal affinity chromatography. Fractions containing purified proteorhodopsin displayed a characteristic pink color, corresponding to an absorption maxima at 520 nm, indicative of retinal complexed with proteorhodopsin (Figure 8B) (30). As can be seen in Figure 8A, the majority of proteorhodopsin eluting off the Ni^{+} column was of high purity (>90%). The predicted molecular weight of proteorhodopsin is 27 kDa, however, purified proteorhodopsin had an apparent molecular mass of approximately 22 kDa (Figure 8A), which is similar to what has been previously reported (31, 32). The final proteorhodopsin yield from wild type *R. sphaeroides* was under 1 mg of purified protein per liter of culture, while the ΔRCLH mutant produced approximately 3 mg of proteorhodopsin per liter of culture. Given the general difficulty associated with

the expression and purification of membrane proteins, our proteorhodopsin yields reveal the utility of our vector for membrane protein production.

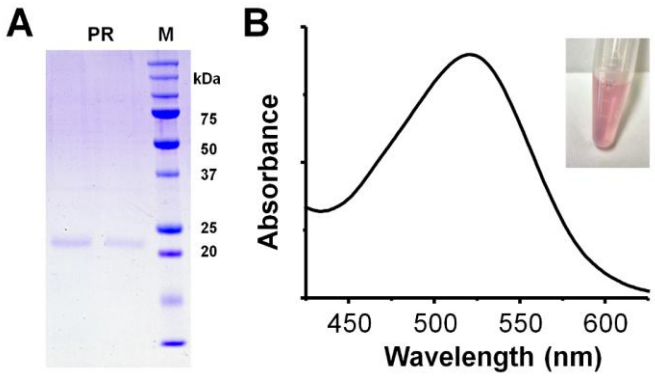

Figure 8

Expression and purification of proteorhodopsin. A Representative SDS gel showing the elution of proteorhodopsin (PR) from the Ni+ affinity column. Proteorhodopsin eluted from the Ni+ column was found to be approximately 22 kDa and of high purity. **B** The eluted fractions exhibited the pink color (insert) and absorbance spectrum with a maximum of 520 nm characteristic for proteorhodopsin.

DISCUSSION

Rhodobacter has the potential to be an excellent platform for biotechnology applications, such as the production of value-added compounds and industrially-relevant membrane proteins. However, for these applications to be practical, a stable expression platform needs to be available. Unlike protein expression systems developed for *E. coli* over the past several decades, few viable examples exist for *Rhodobacter*. We have directly addressed this need by developing and implementing a new, robust and modular, plasmid-based expression system designed specifically for *Rhodobacter*.

Many of the currently available *Rhodobacter* expression systems utilize either the *puf* or the *puc* promoters, which drive the expression of the light harvesting machinery (14, 15, 18, 33). By utilizing the native *puf* promoter, we are able to leverage the intricate regulation mechanisms already present in the cell, similar to what has been done in the

past for P_{lac} in *E. coli*. For example, transcription through the *puf* region is regulated by both oxygen and light through the PpsR and AppA repressor and antirepressor system as well as the redox-sensing PrrBA system (29, 34, 35). PpsR is a transcriptional repressor, which regulates the transcription of proteins involved in carotenoid and bacteriochlorophyll biosynthesis, as well as both the *puf* and *puc* operons (29, 36). The interaction of AppA with PpsR removes the complex from the chromosome, thus alleviating transcriptional repression. The utilization of a repressor and antirepressor in the regulation of P_{puf} presents parallels to the regulation of P_{lac} in *E. coli*. We believe that our vectors present a perfect starting point for making P_{puf} as ubiquitous in *Rhodobacter* expression systems as P_{lac} is in *E. coli*. Coupling our increasing knowledge regarding the regulation of P_{puf} with standardized and modular vectors such as the ones created in this study, holds a great deal of potential for engineering advanced expression systems to suit future experimental needs. Furthermore, recent work has also shown that the AppA-PspR system is also regulated by blue light (29) and this provides yet another way in which our expression system might be further optimized. It would be especially interesting in the future to fine-tune the P_{puf} response in our novel vectors to blue light and create the first truly light-inducible system for *R. sphaeroides*.

Ultimately, the combination of novel *R. sphaeroides* strains optimized for the reconstruction of biosynthetic pathways and optimized expression vectors, will create a powerful genetic toolkit for engineering in *Rhodobacter*, much like what is now being done with *E. coli*. The benefits of this sort of application-specific strain and plasmid engineering are well documented (37) and are corroborated by the results of this study,

which demonstrated consistently higher protein yields when utilizing *R. sphaeroides* ΔRCLH as opposed to wild-type *R. sphaeroides*. In addition to the regulation on the transcription level, it is thought the products of the *puf* operon are also regulated on a translational level (20, 38, 39).

A number of reports have demonstrated that organisms reduce translation of open-reading frames through the use of a less efficient RBS, which reduces ribosome loading (40, 41). Secondary structures of mRNA are another common means to retard the passage of the ribosome through the transcript and therefore lowering translation efficiency (40, 42). Such a mechanism could be responsible for the lower levels of DsRed present in strains carrying the vector with the $P_{puf1\text{-}842}$ segment compared to strains with either the $P_{puf1\text{-}1200}$ or $P_{puf843\text{-}1200}$ segments. To ascertain if this was indeed the case, we first analyzed the region directly upstream of the start codon of DsRed to determine if a significant variation in the RBS sequence or its location existed between the three constructs. In the $P_{puf1\text{-}1200}$ and $P_{puf843\text{-}1200}$ segments, a GGAGGA sequence precedes the ATG of the start codon by ten base pairs, very much in line with what has been reported for optimal RBS position for prokaryotic protein expression (40, 41, 43). For the $P_{puf1\text{-}842}$ sequence, an A-G rich motif is also present upstream of the engineered start codon, however, it is located further upstream, by as much as ten additional bases. This is important because the distance between the RBS and the translational start codon is known to be a crucial factor determining the efficiency of protein synthesis (40, 41). It is therefore likely that the difference in RBS placement in our vectors relative to the start codon of DsRed accounts for the difference in the observed levels of reporter production.

By leveraging these differences in the proximity of RBSs to the start codon, we have created a set of plasmids able to drive protein expression at either a high or a moderate level. The ability to choose between a strong and a moderate protein expression level is highly advantageous for pathway engineering as it allows tuning the final amount of expressed protein without adjusting growth conditions.

In addition to the high levels of expression of soluble proteins, as demonstrated by DsRed, the newly developed vectors successfully produced high yields of proteorhodopsin, a model membrane protein. The final yield of pure proteorhodopsin from the ΔRCLH deletion strain (approximately 3 mg/L) is comparable with what has been demonstrated for the same protein in *E. coli* (31) and clearly displays the utility of our system for the production of membrane proteins. However, in order to achieve these levels of proteorhodopsin production in *E. coli*, extensive optimization of promoter strength, growth medium and growth conditions were required (31). The fact that we were able to achieve high proteorhodopsin yields with no optimization underscores the simplicity afforded by utilizing our newly developed plasmid system for expression and purification of membrane proteins from *R. sphaeroides*, without the time and expenses of optimizing a myriad of complex variables.

The newly developed set of vectors presented in this study represents a simple, modular system which researchers can easily employ to clone genes of interest and create multi-enzyme, designer circuits utilizing a single plasmid. Depending on which *puf* promoter section is employed, strong or moderate expression can be achieved, giving researchers in the ability to modulate the protein expression levels, including those of

recalcitrant membrane proteins. In addition to the choice of promoters, this expression system is further tunable based on the oxygen concentration in the medium as was demonstrated in Figure 2. From a bioengineering perspective, this is exciting not only because potential toxicity can be mediated by regulating the amount of protein produced, but also because it allows for regulation of individual enzymes when working with complex multi-enzyme pathways.

As demonstrated by the overexpression of proteorhodopsin, these vectors are especially useful for the production of heterologous and/or problematic membrane proteins. Although *R. sphaeroides* has long been a model organism for expression of membrane proteins, our results provide a simple, standardized and modular way to build and routinely express membrane proteins. The BioBrick™ design utilized herein is also a great tool for researchers looking to characterize novel promoters in *Rhodobacter* due to the simple and standardized cloning site, as well as a conveniently placed reporter. A fluorescent protein as a reporter coupled with flow cytometry creates a more sensitive real time display of population dynamics compared to conventional reporters such as beta-galactosidase, which only display the average state of the population. Given that without optimizing the growth conditions we were able to purify quantities of proteorhodopsin similar to those purified out of *E. coli* after significant optimization, it is likely that with optimized growth conditions and perhaps strain evolution, the levels of pure protein isolated from these cultures could be pushed quite a bit higher, thus surpassing the amount of proteorhodopsin produced in any other recombinant host. Moreover, it seems plausible that these optimizations would translate well towards the

expression of other membrane proteins, as well. Taken together, these results provide researchers with a toolkit, which greatly enhances our ability to optimize metabolic flux and ensure the highest efficiencies and product yield, instead of more common; one size fits all expression systems.

Abbreviations: OG n-Octyl-β-D-glucopyranoside, LDAO N,N-Dimethyldodecylamine N-Oxide, DMSO dimethyl sulfoxide.

Chapter 3. A tale of two reductases: extending the bacteriochlorophyll biosynthetic pathway in *E. coli*

This chapter is adopted from: A tale of two reductases: extending the bactriochlorophyll biosynthetic pathway in *E. coli*. Ilya B. Tikh, Maureen B. Quin and Claudia Schmidt-Dannert. Submitted for publication in PLoS One.

Chapter 3 Overview

The creation of a synthetic microbe that can harvest energy from sunlight to drive its metabolic processes is an attractive approach to the economically viable biosynthetic production of target compounds. Our aim is to design and engineer a genetically tractable non-photosynthetic microbe to produce light-harvesting molecules. Previously we created a modular, multienzyme system for the heterologous production of intermediates of the bacteriochlorophyll (BChl) pathway in *E. coli*. In this report we extend this pathway to include a substrate promiscuous 8-vinyl reductase that can accept multiple intermediates of BChl biosynthesis. We present an informative comparative analysis of homologues of 8-vinyl reductase from the model photosynthetic organisms *Rhodobacter sphaeroides* and *Chlorobaculum tepidum*. The first purification of the enzymes leads to their detailed biochemical and biophysical characterization. The data obtained reveal that the two 8-vinyl reductases are substrate promiscuous, capable of reducing the C8-vinyl group of Mg protoporphyrin IX, Mg protoporphyrin IX methylester, and divinyl protochlorophyllide. However, activity is dependent upon the presence of chelated Mg in the porphyrin ring, with no activity against non-Mg chelated intermediates observed. Additionally, CD analyses reveal that the two 8-vinyl reductases appear to bind the same substrate in a different fashion. Furthermore, we determined that

the different rates of reaction of the two 8-vinyl reductases both *in vitro*, and *in vivo* as part of our engineered system, results in the suitability of only one of the homologues for our BChl pathway in *E. coli*. Our results offer the first insights into the different functionalities of homologous 8-vinyl reductases. This study also takes us one step closer to the creation of a non-photosynthetic microbe that is capable of harvesting energy from sunlight for the biosynthesis of molecules of choice.

INTRODUCTION

Sunlight is an abundant and sustainable energy source that is captured by photosynthetic organisms and converted into chemical energy for growth and survival. Utilization of the photosynthetic machineries of light harvesting organisms plays an important role in the bioproduction of fuels and chemicals (1-4). Engineering light capture and conversion into genetically tractable, non-photosynthetic and robust microorganisms already used for industrial processes represents an alternative approach (5,6). Such designer microbes could be engineered to synthesize a range of valuable and novel compounds from inexpensive carbon sources where light-energy drives otherwise expensive synthetic reactions (7).

The first steps on the path towards engineering a non-photosynthetic microorganism able to harvest light-energy are to install either simple light-driven proton pumps (8) or more powerful photosynthetic reaction centers (5,6). Both systems require functional assembly of a biosynthetic pathway for carotenoid-derived pigments, and reaction centers also require (bacterio)chlorophyll ((B)Chl) pigments for function. While

engineering of carotenoid pathways into various hosts has been shown (9,10), complete reconstruction of a BChl biosynthetic pathway remains to be demonstrated, a formidable task owing to the complexities of the reaction pathway and enzymes involved (Figure 9). As a first step towards this goal, we created a modular system for the high level production of porphyrins, including protoporphyrin IX (P^{IX}), by assembling genes involved in heme biosynthesis (HemA-F) in *E. coli* (11). P^{IX} is the common intermediate between the heme and BChl biosynthetic pathways (12-14) and is committed to Bchl biosynthesis upon insertion of a central Mg^{2+} catalyzed by a multi-subunit magnesium (Mg-) chelatase enzyme complex BchHID (homologues of H, namely S and T, are present in some bacteria like the green sulfur bacterium *Chlorobaculum tepidum*) (15). The chelatase subunit BchH interacts with the SAM-dependent methyltransferase BchM, which methylates MgP^{IX} at the C13-carboxyl group, resulting in MgP^{IX} monomethyl ester ($MgP^{IX}ME$) (16-18). Co-expression of BchSID and BchM from *C. tepidum* in our P^{IX} overproducing *E. coli* strain resulted in high level production of P^{IX}, $P^{IX}ME$, MgP^{IX} and $MgP^{IX}ME$ (15). Detailed *in vitro* studies provided insights into enzyme interactions and kinetics and revealed that BchM also methylates P^{IX}, resulting in the accumulation of the "dead-end" product $P^{IX}ME$, which cannot be chelated by BchSID (15).

Following chelation and methylation of P^{IX}, the characteristic fifth ring of the chlorin molecule is formed under anaerobic conditions by the radical-SAM cyclase BchE, or under aerobic conditions by AcsF, producing divinyl protochlorophyllide (DVP) (17,19,20). Reduction of the D pyrrole ring of DVP to produce chlorophyllide is either catalyzed by a light-independent, nitrogenase like (DPOR) or by a light-dependent

(LPOR) protochlorophyllide reductase (21-26). An NADPH-dependent reduction of the C8-vinyl group to an ethyl group by 8-vinyl reductase BciA results in chlorophyllide *a* (27-30). Beyond this, another seven enzymes are required to complete the biosynthesis of bacteriochlorophyll *a* (14). Figure 9 shows the upper part of the BChl pathway; depending on the substrate specifities of the biosynthetic enzymes, the order in which they operate may differ from the sequence shown.

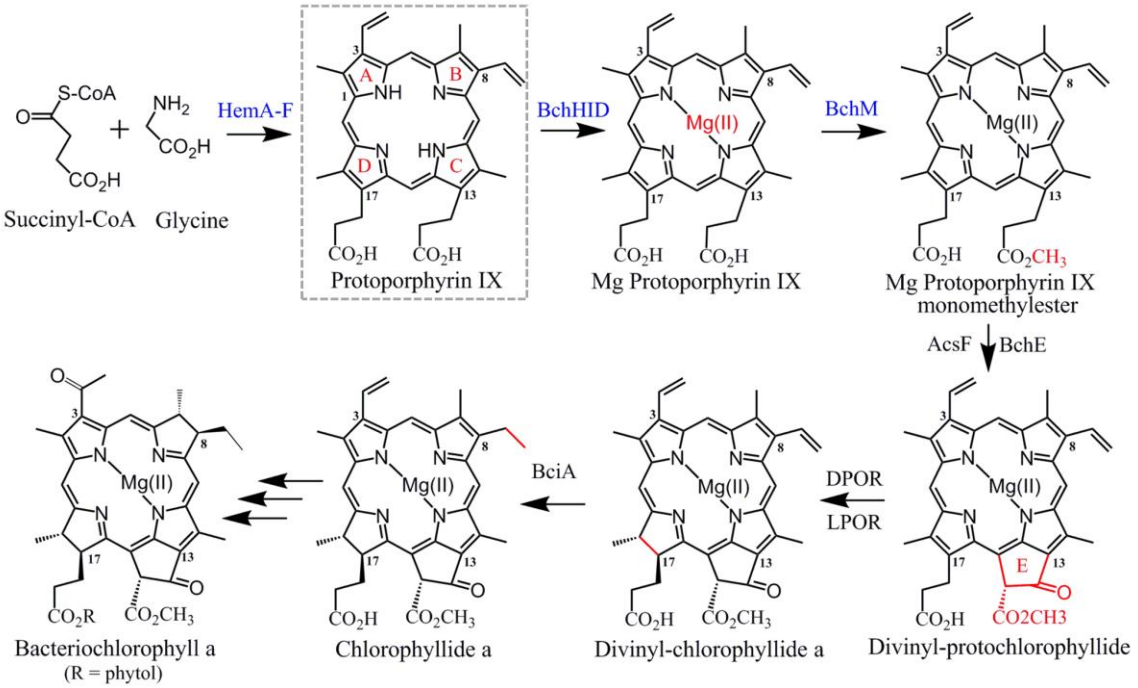

Figure 9.

Engineered pathway design for the heterologous production of BChl. Using succinyl-CoA and glycine as precursor molecules, expression of the heme pathway enzymes HemA-F in *E. coli* results in production of P^{IX} as common intermediate of the heme and BChl biosynthetic pathways. Addition of the BChl enzymes magnesium chelatase (BchHID) and methyltransferase (BchM) yields MgP^{IX} and $MgP^{IX}ME$ in *E. coli* (11,15). Subsequent steps have not yet been functionally assembled in a heterologous system and depending on the enzymes substrate specificities, the order in which the enzymes operate may differ from the depicted pathway. Briefly, formation of the characteristic fifth E ring of chlorophylls is catalyzed by two unrelated and yet to be biochemically characterized cyclases AcsF (aerobic) (20) or BchE (anaerobic) (19). The D pyrrole ring is reduced either by a light-dependent, nitrogenase-like (LPOR, three-subunit enzyme BchLNB) or a light-independent (DPOR) protochlorophyllide reductase; both enzyme have been biochemically characterized (23-26). Reduction of the C8-vinyl group of BChl intermediates is catalyzed by the NADPH-dependent reductase BciA (27,65) investigated in this study. Seven additional enzymatic steps are required for production of Bchl *a* (14).

While some of these additional enzymes have been functionally expressed and biochemically characterized *in vitro* (e.g. LPOR and DPOR (23,25,31)), other steps of this complex pathway have only been elucidated by gene knockouts/deletions, complementation, and mutational studies (14,32,33). Many of these enzymes form complexes, catalyze novel reactions and may interact with yet to be identified protein partners (34), making biochemical studies as well as heterologous pathway reconstitution particularly challenging. In our quest towards recombinant BChl biosynthesis, we report the extension of the BChl biosynthetic pathway in *E. coli* to include a 8-vinyl reductase. Recent studies have indicated that various homologues of the 8-vinyl reductase BciA are substrate promiscuous *in vivo* and can reduce the C8-vinyl group of different intermediates of the BChl pathway (35,36). We hypothesized that including a 8-vinyl reductase as the next step in our pathway would result in reduction of the C8-vinyl group of multiple BChl intermediates that do not have the fifth ring of divinyl protocholorophyllide (DVP) (Figure 9), thereby possibly removing barriers to the efficient turnover of P^{IX} in our engineered system.

We demonstrate co-expression of the heme biosynthetic pathway in conjunction with *C. tepidum* BchSID and BchM with two separate homologues of BciA from *Rhodobacter sphaeroides* (*RS*BciA) (30) and *C. tepidum* (*CT*BciA) (27). We discovered that while *CT*BciA is capable of reducing the C8-vinyl group of several different intermediates in the BChl pathway, *RS*BciA is surprisingly completely inactive in our recombinant system. We therefore conducted a full purification and *in vitro* characterization of the two BciA homologues to elucidate possible mechanisms for their

different activities. Results show that both *RS*BciA and *CT*BciA are substrate promiscuous *in vitro*, however, the two enzymes exhibit very different catalytic turnover efficiencies. Biophysical characterization suggests that these differences may be related to different mechanisms of substrate binding. This study provides useful insights for BChl pathway design and another enzymatic step in the complex pathways leading to (B)Chls.

MATERIALS AND METHODS

Materials

All chemicals were obtained from Sigma Aldrich (St. Louis, MO), unless otherwise stated. Restriction enzymes and DNA polymerases were purchased from New England Biolabs (Ipswich, MA) and were used according to manufacturers' procedures. Protein ladder was purchased from Biorad (Hercules, CA).

Bacterial strains, plasmids and growth conditions

All bacterial cultures were grown under aerobic conditions at 37 °C with shaking at 220 rpm in Luria-Bertani (LB) medium supplemented with chloramphenicol (50 µg/ml), kanamycin (30 µg/ml), ampicillin (100 µg/ml), and streptomycin (50 µg/ml) as required for plasmid maintenance. The pET30a (+) vector was purchased from EMD Millipore (Billerica, MA). *R. sphaeroides* 2.4.1 and *C. tepidum* TLS were acquired from the ATCC collection (Manassas, VA). *E. coli* JM109 was used for all genetic manipulations and *E. coli* BL21 (DE3) was used for protein expression.

Construction of plasmids

Expression vectors containing *rsbciA* and *ctbciA* were constructed as follows: pET30-*rsbciA* was constructed by amplifying *bciA* (rsp_3070, GenBank Accession Number: 3721347) from *R. sphaeroides* 2.4.1 genomic DNA using the gene specific primers P1 (forward) and P2 (reverse) (Table 2), with the reverse primer introducing a His_6 tag for purification purposes. Similarly, pET30-*ctbciA* was constructed by amplifying *bciA* (ct_1063, GenBank Accession Number: 1006951) from *C. tepidum* TLS genomic DNA using primers P3 and P4 (Table 2), again introducing a His_6 tag for purification. The PCR products were digested with *Nde*I and *Not*I and were cloned into pET30a (+) which was digested with the same enzymes, and the sequence of the resulting plasmids was verified by DNA sequencing. Vector pCDFBB-*rsbciA* was constructed by amplifying *rsbciA* using primers P5 and P6, digesting with *Bgl*II and *Not*I and ligating into pCDFBB-GFP which was digested with the same enzymes. Plasmid pCDFBB-*bchM* was created by amplifying *bchM* from pCDF-*bchM* (15) using primers P7 and P8 and the resulting PCR product inserted into the *Bgl*II and *Not*I sites of pCDFBB. Similarly, *ctbciA* was cloned into the *Bgl*II and *Not*I sites of pCDFBB after amplification with P9 and P10, generating pCDFBB-*ctbciA*. Subsequent gene stacking of *rsbciA* and *ctbciA* with *bchM* was performed via standard BioBrick techniques, described elsewhere (37), generating plasmids pCDFBB-*bchM-rsbciA* and pCDFBB-*bchM-ctbchiA*. Similarly, *bchJ* (rsp_0280, GenBank Accession Number: 3719192) was amplified from *R. sphaeroides* 2.4.1 genomic DNA using primers P11 (forward) and P12 (reverse) containing *Bgl*II and *Not*I site, respectively. Following restriction enzyme digest, the PCR product was ligated

into the same sites of pCDFBB to create pCDFBB-*bchJ*. *bchJ* was stacked with *bchM* and *rsbciA* in a pCDFBB backbone following standard BioBrick techniques, as described above. Plasmids used in this study are listed in Table 3.

Table 2. Primers described in Chapter 3.

P1	5'-ATAcatatgTCCGAGACCGCCCCCCTGC-3'
P2	5'-ATgcggccgcTCAATGGTGATGGTGATGGTGACGATTTCCGGGCGATCCTTCTGC-3'
P3	5'-ATAcatatgTCATCTTCGTCTGTACTGGCTG-3'
P4	5'-ATgcggccgcTCAGTGGTAGTGGTAGTGGTAGAACATCGCGTGCGCGCCGAG-3'
P5	5'-ACAGATCTAATGTCCGAGACCGCCCCCCTGC-3'
P6	5'-ATGCGGCCGCTCAATGGTGATGGTGATGGTGACGATTTCCGGGCGATCCTTCTGC-3'
P7	5'-CAAGATCTATGAGCAGCCCATCATTCAAC-3'
P8	5'-ATGCGGCCGCTCAGCGGCGAGCCACG-3'
P9	5'-atagatctATGTCATCTTCGTCTGTACTGGCTG-3'
P10	5'-ATgcggccgcTCAGAACATCGCGTGCGCGCCGAG-3'
P11	5'-ACagatctATGACCGCGCATGACCAGCG-3'
P12	5'-AAgcggccgcTCAGCGCCCCTTGGGCAG-3'

Table 3. Strains and plasmids used in Chapter 3.

Plasmid	**Relevant properties**	**Source**
pACmod-*hemABCD*	Constitutive expression of *R. capsulatus hemA* and *E. coli hemB, hemC,* and *hemD*	(1)
pBBR-*hemEF*	Constitutive expression of *Synechocystis hemE* and *E. coli hemF*	(1)
pUCmod-*bchDIS*	Constitutively expression of *C. tepidum bchD, bchI* and *bchS*	(2)
pCDFBB	Empty pCDFBB (BioBrick^{Tm}) vector	unpublished
pCDFBB-*bchM*	Constitutive expression of *C. tepidum bchM*	This paper
pCDFBB-*rsbciA*	Constitutive expression of *R. sphaeroides bciA*	This paper
pCDFBB-*ctbciA*	Constitutive expression of *C. tepidum bciA*	This paper
pCDFBB-*bchM-rsbciA*	Constitutive expression of *C. tepidum bchM* and *R. sphaeroides bciA*	This paper
pCDFBB-*bchM-ctbciA*	Constitutive expression of *C. tepidum bchM* and *bciA*	This paper
pCDFBB-*bchJ*	Constitutive expression of *R. sphaeroides bchJ*	This paper
pCDFBB-*bchM-bchJ*	Constitutive expression of *C. tepidum bchM* and *R. sphaeroides bchJ*	This paper
pCDFBB-*bchM-rsbciA-bchJ*	Constitutive expression of *C. tepidum bchM* and *R. sphaeroides bciA and bchJ*	This paper

Sequence and phylogenetic analysis

All protein alignments were performed in MEGA5.1 (38) using the ClustalW algorithm (39). The protein sequences were identified using NCBI BLAST (40) with *RS*BciA as the query and hit proteins limited those containing above 30 % sequence identity. The phylogenetic tree was created in MEGA5.1 using the default parameters for the neighbor-joining algorithm (41) with a bootstrap test of phylogeny (500 replicates) (42). Protein modeling was conducted using the automated mode in MODELLER v 9.12 (43) and models were visualized using PyMOL v 1.6 (Schrödinger, LLC).

Biosynthesis of protoporphyrin IX derivatives in *E. coli*

The production of magnesium porphyrins in *E. coli* harboring plasmids pAC-*hemAD*, pBBRB-*hemEF*, pUCMOD-*SID* and pCDF-*bchM* has been described previously (11,15). Briefly, *E. coli* were transformed with the plasmids and were incubated at 30 °C shaking at 220 rpm in 4 ml of LB broth supplemented with chloramphenicol (50 µg/ml), kanamycin (30 µg/ml), ampicillin (100 µg/ml), streptomycin (50 µg/ml) and 1 mM $MgCl_2$ for 48 hours. Porphyrins were extracted from the cells as follows: 0.25 ml of culture was centrifuged at 21000 x *g* for 1 minute and the pellet was resuspended in 1 ml of a water:acetone:methanol mix at a 1:7:2 ratio. The cells were lysed and the pigments were extracted by vortexing for 20 seconds every 10 minutes over the course of two hours. Subsequently, the samples were centrifuged at 21000 x *g* for 5 minutes to remove cell debris and the remaining supernatant was analyzed with an Agilent 1100 HPLC system equipped with an photodiode array detector as previously described (15). Comparison of samples to the integrated peak areas of known concentrations of authentic

P^{IX} and MgP^{IX} standards (Frontier Scientific, Logan, UT) was used to determine porphyrin concentrations. For determination of *in vivo* activity of BciA and BchJ, pigments were extracted from *E. coli* cultures as described above, with pCDFBB-*bchM* being replaced with a pCDFBB vector containing one of the BciA homologues, or BchJ.

Protein expression and purification

E. coli cells were transformed with either pET30-*ctbciA* or pET30-*rsbciA*, and single colonies were used to inoculate 1 L LB supplemented with kanamycin (30 µg/ml). Protein expression was induced at an OD_{600} of 0.6 upon addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. The culture was incubated for an additional 16 hours at 30 °C and cells were harvested by centrifugation. The cell pellet was resuspended in Buffer A (50 mM Tris-HCl, 100 mM NaCl, 5 mM imidazole, pH 8.0) and cells were lysed by sonication. The supernatant was clarified by centrifugation at 15 000 x *g* for 20 minutes and the soluble portion was loaded onto a 5 ml HisTrap™ FF column (GE Healthcare, Piscataway, NJ) which was equilibrated with Buffer A. The protein was eluted from the column over a linear gradient to 100 % Buffer B (50 mM Tris-HCl, 100 mM NaCl, 250 mM imidazole, pH 8.0) over 20 column volumes. The resulting fractions were analyzed by SDS-PAGE to assess protein purity. Fractions determined to be > 95% pure were pooled and the protein concentration was determined by Coomassie Plus Bradford assay (Thermo Scientific, Rockford, IL).

Size exclusion chromatography

Purified BciA was dialyzed overnight into Buffer C (50 mM Tris-HCl, 10 % glycerol, 1 mM DTT, pH 8.0) using 3 kDa cutoff dialysis tubing. Size exclusion chromatography

was performed at a flow rate of 0.5 mL/min by passing 30 µM BciA over a Superdex S200 10/300 GL column (GE Healthcare, Piscataway, NJ) equilibrated with Buffer C. Standard curves were prepared using standards of known molecular weights (Biorad, Hercules, CA).

Circular dichroism analysis

Prior to structural analysis by CD, purified BciA was dialyzed into Buffer D (10 mM NaHPO4, 100 mM NaCl, pH 7.4). The secondary structure of BciA was analyzed using a J-185 spectrometer (JASCO, Easton, MD). Data were collected using 1 nm wavelength intervals from 200 – 260 nm using 7.5 µM of protein in a 1 mm pathlength cuvette. Conformational changes in metal porphyrins were monitored in 1 nm steps between the wavelengths of 350 – 500 nm upon addition of 20 µM metal porphyrin to BciA. Analysis of protein secondary structure content was performed with CDPro using the CONTIN method and SMP56 reference set (44).

8-vinyl reductase enzyme assay

DVP (divinyl protochlorophyllide) was purchased from ChromaDex (Irving, CA). P^{IX}ME and Mg-P^{IX}ME were purified from *E. coli* cells as previously described (15). Enzyme assays were carried out in Buffer E (50 mM Tris-HCl, pH 8.0). Final reaction volumes of 50 µL included 500 µM of either NADH or NADPH as a cofactor, substrate at the concentrations listed below, and 1 µM of purified BciA. In the case of control reactions, an equal volume of Buffer E was added in place of BciA. The reactions were incubated for 30 minutes at 37 °C for DVP (2 µM) or 18 hours at 25 °C for other porphyrin substrates (75 µM). Reactions were quenched by addition of 70 µL acetone.

Precipitated protein was removed by centrifugation. The reaction products were analyzed on an Agilent 1100 HPLC system equipped with an photodiode array detector. Porphyrin derivatives were separated using a 25 mm Zorbax C-18 column (Agilent Technologies, Santa Clara, CA) using a gradient of 80 % - 100 % methanol over 55 min at a flow rate of 1 ml/min, using an aqueous buffer (0.1 M ammonium acetate, pH 5.1) for reactions containing a mixture of porphyrins. DVP was resolved using isocratic conditions of 80 % methanol and 20 % aqueous buffer. For structural analysis, mass fragmentation spectra were monitored on a LCQ mass spectrophotometer equipped with an electron spray ionization (ESI) interface (Thermo Finnigan, USA) operating in the positive mode.

RESULTS

Selecting BciA homologues for pathway assembly

Our previously engineered, modular heme biosynthetic pathway in *E. coli* was assembled from *hemA-F* genes selected from diverse bacterial sources (*E. coli*, *B. subtilis*, *R. capsulatus* and *Synechocystis)* based on available biochemical data and reported activities (11). We then chose to extend the heme pathway with the first two steps of the BChl pathway comprised of BchHID(S,T) and BchM from the green sulfur bacterium *C. tepidum* (15). Green bacteria like *C. tepidum* are unique in that they are able to produce different types of Chls and Bchls, and encode in their genomes several homologs (BchS, T) of the large subunit (BchH) of the magnesium chelatase, which may play a role in regulating the types of (B)Chls produced (see (15) for a comprehensive

discussion). These are useful properties that may later become important for heterologous production of different (B)Chl structures.

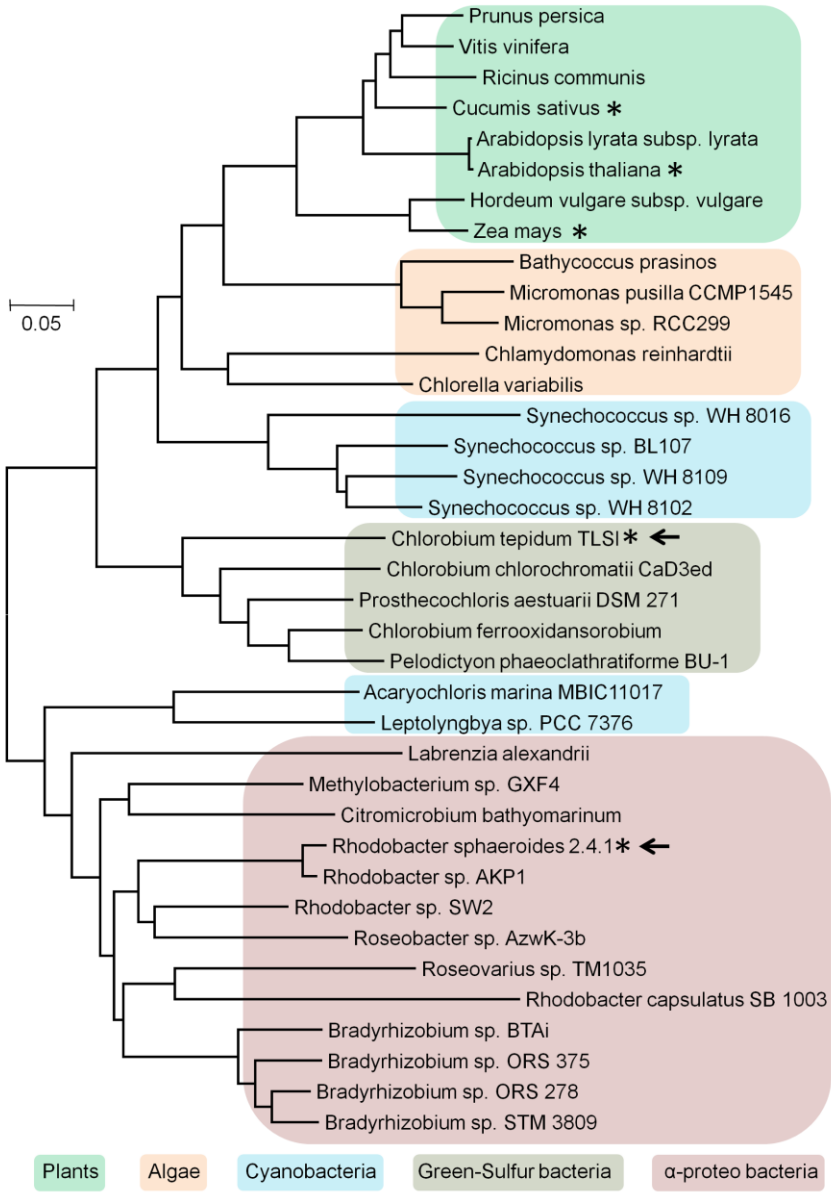

Figure 10.

Phylogenetic analyses of 8-vinyl reductases to select candidates for pathway engineering. BLAST searches using *RS*BciA as search template identify 37 putative 8-vinyl reductases that share greater than 30 % sequence identity. Homologues cluster with other members of the various kingdoms of life, highlighted by colored boxes. Five of the identified 8-vinyl reductases have been previously characterized, highlighted with asterisks. BciA from *R. sphaeroides* and *Chlorobaculum* (*Chlorobium*) *tepidum*, marked with arrows, are the two bacterial characterized 8-vinyl reductases that were selected to extend our engineered BChl pathway in *E. coli*.

To apply the same combinatorial rationale to the current BChl pathway extension, we searched all kingdoms of life for suitable 8-vinyl reductase candidates. Using the recently described 8-vinyl reductase BciA from the well characterized model photosynthetic purple bacterium *R. sphaeroides* (30,45) to guide our searches, we identified 38 putative BciA-like 8-vinyl reductases in photosynthetic organisms belonging to five kingdoms (Figure 10). Of those homologues identified, only five have been previously described in literature (27,28,30,36,46), and activity has been shown either by gene knockouts in the native host (*R. sphaeroides, C. sativus, A. thaliana, Z. mays*), or by heterologous expression and analysis of cell lysate constituents (*C. tepidum, A. thaliana*). However, no biochemical data exists to provide insights into the kinetics or mechanistic details of 8-vinyl reductase, which could be used to select the best enzyme for our pathway. Therefore, for comparative purposes we selected two BciA homologues from *R. sphaeroides* and *C. tepidum* for BChl pathway extension based upon information gathered from sequence alignments.

The selected *R. sphaeroides* and *C. tepidum* BciA homologues are the most closely related of the previously characterized bacterial enzymes, sharing 53 % sequence identity, and were therefore hypothesized to behave in a similar fashion. Additionally, both *RS*BciA and *CT*BciA have a well-defined GxxGxxG motif (Rossmann-fold, PF13460) (47) for binding of NAD(P)H which is essential for 8-vinyl reductase activity (29) (Figure 11). Attempts to model the 3D structure of the two full-length proteins using MODELLER v 9.12 (43) did not produce reliable results as no template structure with sufficient sequence similarity could be identified. However, the N-termini (residues 1-

220) of both BciA homologues could be modeled using the Rossmann fold containing biliverdin IX β reductase as a template (PDB: 1HDO) (48), corroborating the expectation that both enzymes should carry out an NAD(P)H-dependent reduction. Finally, we expected that the similarity of these genes to other components of our existing pathway, which includes members already derived from the closely related *R. capsulatus* and *C. tepidum*, would facilitate expression in *E. coli*.

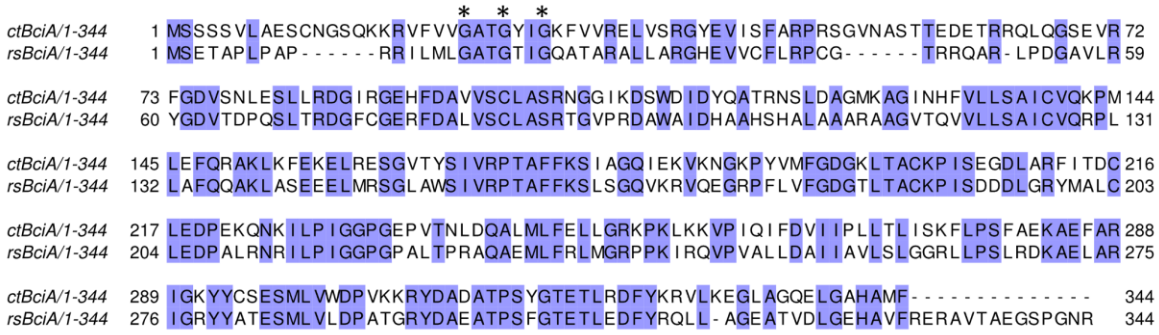

Figure 11.

Sequence alignment of *Chlorobaculum tepidum CT*BciA and *Rhodobacter sphaeroides RS*BciA. The two 8-vinyl reductases share 53 % sequence identity. Conserved residues are highlighted in blue. The conserved GxxGxxG motif, required for NAD(P)H binding (47), is marked with asterisks.

8-vinyl reductase activity in *E. coli* cells expressing HemA-F, BchSID and BchM

For expression of the extended BChl pathway in *E. coli*, genes encoding *RS*BciA and *CT*BciA were cloned into our in house pCDFBB plasmid (37). The BioBrick plasmid system has been designed for the straightforward stacking of several genes thereby facilitating pathway engineering, and constitutive expression is driven by a modified *lac* promoter. The pCDFBB-*ctbciA* and pCDFBB-*rsbciA* plasmids used in this study are compatible with the previously created pAC-*hemAD*, pBBR-*hemEF*, and pUCMOD-*SID* plasmids, allowing expression of the entire 11 gene pathway in a heterologous host. Additionally, by stacking *rsbciA* and *ctbciA* separately onto the

pCDFBB-*bchM* plasmid, we maintained our original four plasmid system, reducing the likelihood of excessive metabolic burden placed on the cells due to different selective markers.

Using this plasmid system, we previously showed that P^{IX} overproducing *E. coli* expressing the Mg-chelatase BchSID complex produced MgP^{IX}. We also showed that pathway extension with methyltransferase BchM resulted in the production of both $MgP^{IX}ME$ and the "dead-end" product $P^{IX}ME$ (15), findings which we also noted in this study (Table 4). However, in our current extended Bchl pathway, P^{IX} overproducing cells expressing BchSID and *CT*BciA alone produced two new compounds, mono-vinyl P^{IX} (mvP^{IX}) and mono-vinyl MgP^{IX} ($mvMgP^{IX}$), indicating that *CT*BciA was capable of reducing the C8-vinyl group on both the Mg chelated and the non-Mg chelated porphyrin molecule. Furthermore, upon expression of BchSID with BchM and *CT*BciA, $MgP^{IX}ME$ was no longer observed, with the concurrent appearance of the new product mono-vinyl $MgP^{IX}ME$ ($mvMgP^{IX}ME$). While it is not known whether the methyltransferase BchM is producing less $MgP^{IX}ME$ when expressed in combination with the 8-vinyl reductase, it would appear that *CT*BciA catalyzes full conversion of the divinyl form of $MgP^{IX}ME$ to the mono-vinyl form. Additionally, levels of the "dead-end" product $P^{IX}ME$ were reduced upon coexpression with *CT*BciA, and mono-vinyl $P^{IX}ME$ ($mvP^{IX}ME$) was produced. Notably, upon addition of *CT*BciA, there is a significant shift in the comparative levels of pathway intermediates derived from P^{IX} flowing towards dead-end non-Mg chelated products ($P^{IX}ME$ and $mvP^{IX}ME$) or towards Mg-chelated products (MgP^{IX}, $MgP^{IX}ME$, $mvMgP^{IX}$ and $mvMgP^{IX}ME$). Without the 8-vinyl reductase, the

BchSID-BchM extended pathway has a preference to produce non-Mg chelated products over Mg-chelated products at a ratio of 2:1. With 8-vinyl reductase, this ratio is shifted to 1:1, improving the pathway balance and resulting in more efficient use of the precursor P^{IX}. Therefore, *CT*BciA can reduce the C8-vinyl group on a diverse range of intermediates in the Bchl pathway.

Table 4. *In vivo* reduction of BChl pathway intermediates.
E. coli cells expressing HemA-F and the magnesium chelatase complex BchSID produce P^{IX} and MgP^{IX}. Addition of the methyl transferase BchM results in production of both P^{IX} ME and MgP^{IX} ME. Expression with the 8-vinyl reductase *CT*BciA in the presence and absence of BchM leads to the production of mono-vinyl forms of pathway intermediates. *RS*BciA is not active in our *in vivo* system. Abbreviations: P^{IX} - protoporphyrin IX, MgP^{IX} - Mg-protoporphyrin IX, P^{IX}ME - protoporphyrin IX methylester, MgP^{IX}ME - Mg-protoporphyrin IX methylester, mvP^{IX} - mono-vinyl protoporphyrin IX, mvMgPIX - mono-vinyl Mg-protoporphyrin IX, mvP^{IX}ME - mono-vinyl protoporphyrin IX methylester, $mvMgP^{IX}$ME - mono-vinyl Mg-protoporphyrin IX methylester.

	Accumulation of BChl pathway intermediates [mg L^{-1}]							
	P^{IX}	**MgP^{IX}**	**P^{IX}ME**	**MgP^{IX}ME**	**mvP^{IX}**	**$mvMgP^{IX}$**	**mvP^{IX}ME**	**$mvMgP^{IX}$ME**
HemA-F+BchSID	80.9 ± 17.2	45.8 ± 7.5	ND	ND	ND	ND	ND	ND
BchSID BchM	76.6 ± 24.1	31.2 ± 7	17.2 ± 5	9.1 ± 3.4	ND	ND	ND	ND
BchSID RSBciA	90.2 ± 2.2	28.4 ± 4.6	ND	ND	ND	ND	ND	ND
BchSID CTBciA	75.9 ± 13.6	25.4 ± 5.5	ND	ND	5.8 ± 1.4	7.5 ± 3.9	ND	ND
BchSID RSBciA BchM	74.7 ± 25.4	37.6 ± 2	14.6 ± 2.6	9.5 ± 1	ND	ND	ND	ND
BchSID CTBciA BchM	68.9 ± 6.3	34.4 ± 10.1	9.1 ± 1.3	ND	4.3 ± 0.4	5.2 ± 3.1	2.9 ± 0.2	3.4 ± 0.6

Surprisingly, we found that *RS*BciA does not behave in a similar fashion to *CT*BciA. We did not detect mono-vinyl forms of any of the pathway intermediates upon coexpression of this 8-vinyl reductase with BchSID and BchM (Table 4), suggesting that *RS*BciA is either inactive in *E. coli* or does not display the same substrate promiscuity as *CT*BciA and instead specifically acts on divinyl protochlorophyllide (DVP) (refer to

Figure 9 for structures). This was an unexpected finding, as our BChl pathway design approach had previously been successful, enabling construction of pathways using enzymes from diverse bacterial sources (11,15). Our sequence analyses had given no prior indication that *RS*BciA would be different from *CT*BciA and not a suitable candidate for BChl pathway engineering.

We considered the possibility that another 8-vinyl reductase from *R. sphaeroides*, BchJ (49), may have been more suitable for heterologous expression in *E. coli*. To explore the potential activity of BchJ, we cloned *rsp_0280* from *R. sphaeroides* and expressed the protein (*RS*BchJ) in a similar manner to the BciA homologues. However, no mono-vinyl intermediates were observed in cultures expressing *RS*BchJ with BchSID and BchM. Previously, it had been suggested that BchJ plays a substrate channeling role, rather than acting as a 8-vinyl reductase (27), therefore we hypothesized that BchJ could be involved in activating *RS*BciA. To test this possibility, we expressed *RS*BchJ in the presence of BchSID, BchM, and *RS*BciA. Once again, no 8-vinyl reductase activity was observed. Furthermore, coexpression of *RS*BchJ with *CT*BciA had no effect on 8-vinyl reductase activity. We therefore excluded BchJ from further analyses. To provide an explanation for the differences in activities of the two 8-vinyl reductase homologues, we set out to purify and characterize reduction of DVP and other P^{IX} derivatives by *RS*BciA and *CT*BciA *in vitro* to inform the design of current and future engineered pathways.

Purification of the two 8-vinyl reductases

To understand the different activities observed for *RS*BciA and *CT*BciA in *E. coli*, we carried out the first purification of 8-vinyl reductases for detailed *in vitro*

characterization. The genes encoding the homologues were cloned separately into the inducible plasmid pET30a(+) with a C-terminal His_6 tag for purification purposes. The two enzymes were overexpressed in *E. coli* as soluble proteins, and were subsequently purified by metal-affinity chromatography upon elution from the column with 100 – 250 mM imidazole. Protein purity of *RS*BciA and *CT*BciA was estimated as > 95 % by SDS-PAGE analysis following a single purification step.

Spectra for purified *RS*BciA and *CT*BciA revealed a maximum peak absorbance at 260 nm, suggesting that an unknown nucleotide(s) copurifies with the enzymes (Figure 12). Attempts to determine its identity by mass spectrometry, however, were unsuccessful. Nonetheless, the conserved Rossmann fold NAD(P)H binding motif near the N-termini of the enzymes (Figure 11) suggests binding of intercellular NAD(P)H shown for other NAD(P)H dependent proteins (50). Binding of this nucleotide is sufficiently tight that it remains with the proteins following their elution as large complexes/aggregates from a size exclusion column (Figure 13). *In vitro* assays with extracts of *E. coli* cells expressing *C. tepidum* BciA showed that the enzyme uses NADPH as co-factor in the reduction of divinyl protochlorophyllide (DVP) (27). The presence of a tightly bound nucleotide following purification of *RS*BciA and *CT*BciA suggested that both purified proteins should be active *in vitro*.

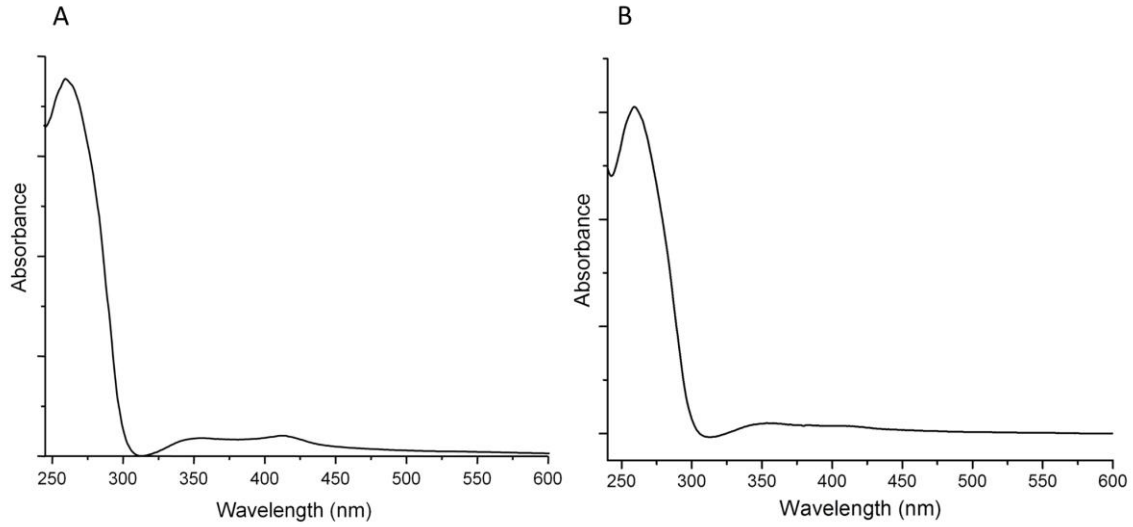

Figure 12.

Nucleotide(s) copurify with recombinant *RS*BciA and *CT*BciA. UV/Vis wavescans of purified 8-vinyl reductases *RS*BciA **(A)** and *CT*BciA **(B)** reveal absorbance maxima at 260 nm. This suggests that an unknown nucleotide(s) co-purifies with the enzymes, likely due to the presence of the conserved NAD(P)H binding site at the N-termini of both proteins.

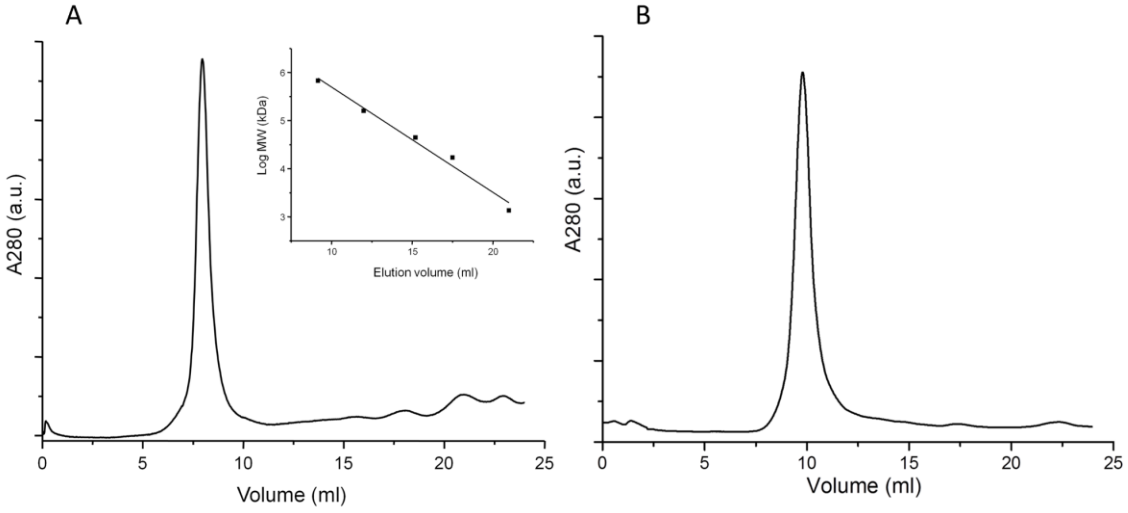

Figure 13.

Recombinant 8-vinyl reductases behave as large complexes/aggregates in solution. (A) *RS*BciA elutes from a size exclusion column after 8.5 mL, close to the void volume of the column. RSBciA appears to be aggregating in solution, despite the presence of the protein stabilizing agent glycerol (10 %). **(B)** *CT*BciA elutes from the same size exclusion column after 10 mL, suggesting that it is forming a large complex close to 600 kDa (as determined using protein standards of known molecular weight).

Reduction of divinyl protochlorophyllide by *RS*BciA and *CT*BciA

To compare catalytic activities of *RS*BciA and *CT*BciA with divinyl protochlorophyllide (DVP), *in vitro* assays with purified protein supplemented with NADPH and DVP were conducted. For both *RS*BciA and *CT*BciA, DVP was reduced to the mono-vinyl form, resulting in the characteristic shift in absorbance maxima from 442 nm to 437 nm (Figure 14) upon reduction of the C8-vinyl group (27,28,51). Identity of the reaction products was further confirmed by LC-MS, revealing the addition of two protons to DVP upon reduction, with a shift in *m/z* from 610 to 612 (Figure 15). These findings confirmed that *RS*BciA is active, albeit not *in vivo* with the engineered pathway. However, we did note a significant difference in the rate of reaction between the two enzymes, with the reaction containing *CT*BciA reaching almost 90 % conversion of divinyl to mono-vinyl after 1.5 hours, and that of *RS*BciA with the same concentration of purified protein only reaching complete conversion after 18 hours (Figure 16), despite numerous attempts at varying reaction conditions (eg. pH, temperature, buffer) Suspecting that *RS*BciA may require an additional cofactor or unknown chaperone, we supplemented the assay with *E. coli* crude cell lysate. Rather than improving reaction efficiency, adding cell lysate decreased the overall conversion of DVP to the mono-vinyl form by *RS*BciA from 100 % to less than 80 % after 18 hours, which could be attributed to an overall lower concentration of *RS*BciA in cell lysate.

It has recently been shown for plant 8-vinyl reductases that these enzymes have broad substrate specificities, with substrate preferences and activities varying according to species (36). In some cases, 8-vinyl reductases are several hundred fold more efficient

at converting a particular intermediate in chlorophyll biosynthesis than another (52). As stated above, we considered the possibility that DVP is not the preferred substrate of *RS*BciA, and that it may be a more specific than *CT*BciA which acts in our engineered *E. coli* cells on any Mg-chelated or unchelated porphyrin IX derivative with a C8 divinyl group (Table 4).

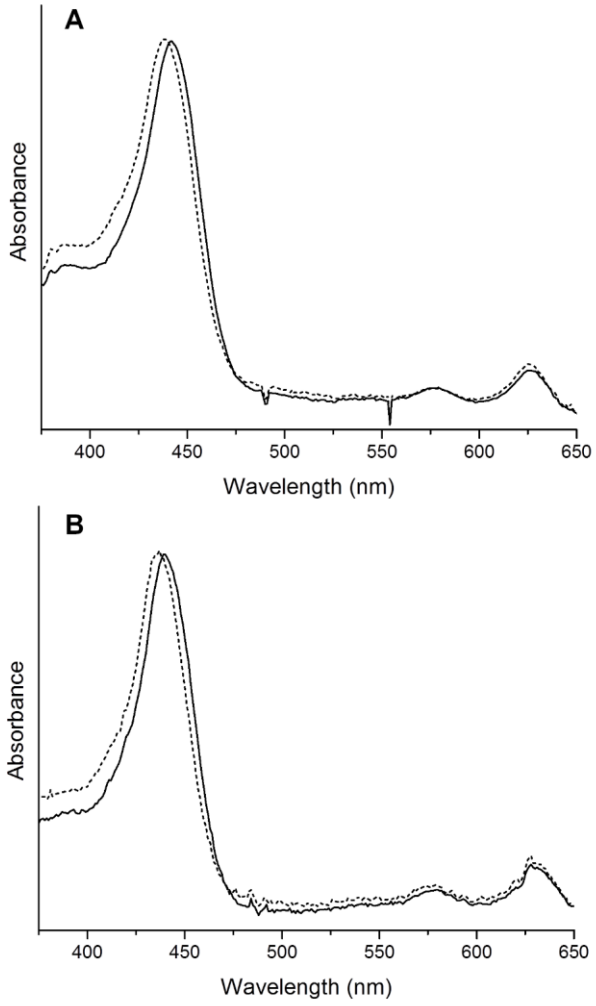

Figure 14.

Reduction of divinyl-protochlorophyllide to mono-vinyl-protochlorophyllide. *In vitro* assays were carried out with NAD(P)H and with purified protein **(A)** *RS*BciA and **(B)** *CT*BciA. Divinyl-protochlorophyllide has a characteristic absorbance maximum of 442 nm (solid line). This shifts 5 nm to 337 nm upon the reduction of the C-8 vinyl group by 8-vinyl reductase (dotted line) (27).

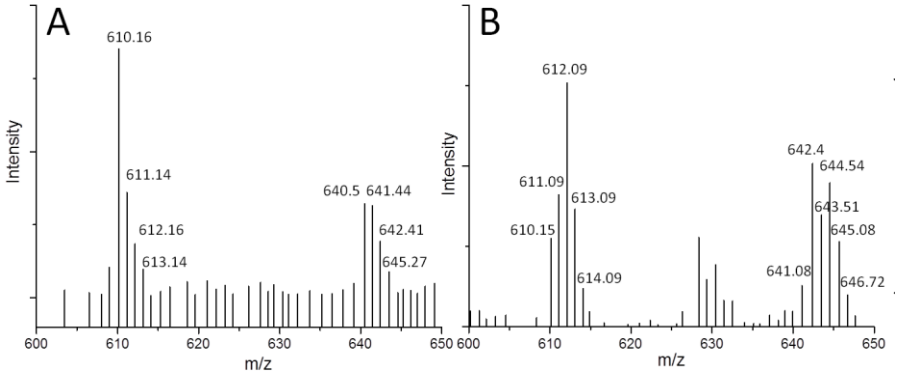

Figure 15.

Mass spectra of divinyl-protochlorophyllide and mono-vinyl-protochlorophyllide. (A) The peak at *m/z* 610 is characteristic of divinyl-protochlorophyllide. **(B)** Upon the reduction of divinyl-protochlorophyllide by *RS*BciA and *CT*BciA to the mono-vinyl form, two protons are added and the mass shifts to *m/z* 612. The peaks at *m/z* 642 and 644, respectively, represent methanol adducts of the two compounds.

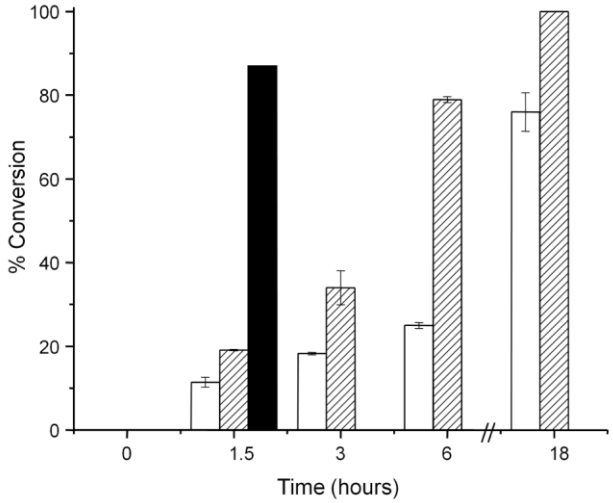

Figure 16.

Reaction efficiency of 8-vinyl reductase with divinyl-protochlorophyllide as substrate. Purified *CT*BciA reduces greater than 85 % DVP to mono-vinyl form in 1.5 hours (black bar). Purified *RS*BciA acts more slowly, reaching 100 % conversion of divinyl to mono-vinyl in 18 hours (hashed bars). Attempts to improve reaction efficiency of *RS*BciA by addition of crude cell lysate to the reaction vessel actually reduced the rate of reaction as well as the overall conversion to less than 80 % in 18 hours (white bars). Error bars are calculated from reactions carried out in duplicate.

Substrate promiscuity of *RS*BciA and *CT*BciA

To determine the substrate preference of *RS*BciA, assays were carried out with purified protein in a reaction mixture containing the cell-extracted BChl pathway

intermediates P^{IX}, MgP^{IX}, $P^{IX}ME$ and $MgP^{IX}ME$, as well as reactions containing commercially available P^{IX} or MgP^{IX}. In all cases, *RS*BciA reduced only the Mg-chelated porphyrins MgP^{IX} and $MgP^{IX}ME$ to the corresponding monovinyl products, no activity was observed with the unchelated P^{IX} derivatives (Figure 17). Additionally, reaction rates were not improved over those obtained with DVP and full conversion of substrates was not observed, suggesting that substrate specificity is not the limiting factor for the inactivity of *RS*BciA in our pathway engineered *E. coli* cells.

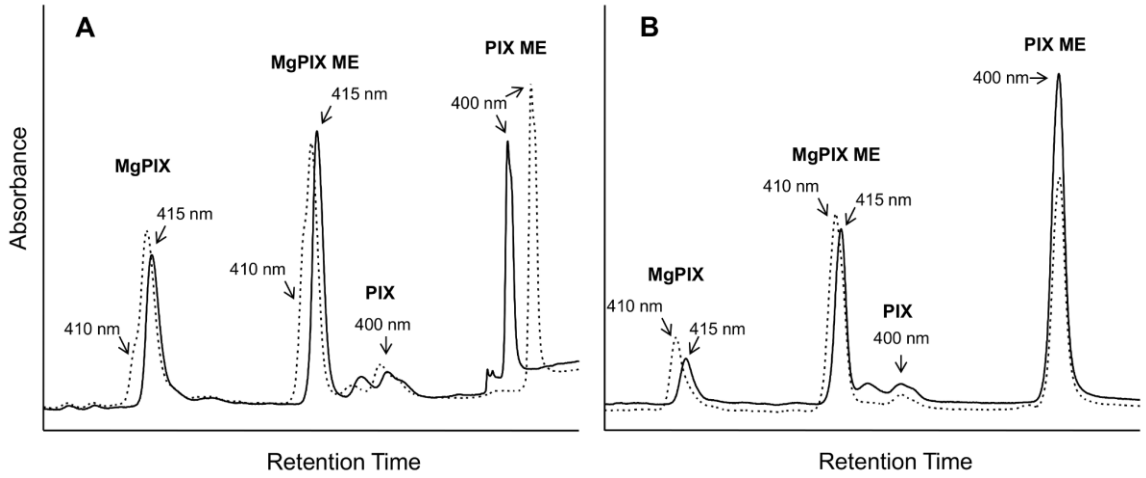

Figure 17.

Substrate promiscuity of purified 8-vinyl reductases with BChl intermediates as determined by shifts in absorbance maxima. Conversion of a mixture of Bchl intermediates (MgP^{IX}, $MgP^{IX}ME$, $P^{IX}ME$) was analyzed by HPLC at a single wavelength (412 nm) to detect all porphyrins present in the reaction mixtures after 18 hours. Reactions with enzyme (dotted traces) and control reactions (solid traces) are shown. Wavelengths displayed above arrows (pointing to peak shoulder or peak maximum) indicate the absorbance maximum measured at that time point, and illustrate the 5 nm absorbance shift which occurs after the reduction of the C-8 vinyl group. (A) Purified *RS*BciA partially reduces the C8-vinyl group of MgP^{IX} and $MgP^{IX}ME$ to generate a peak shoulder for each substrate at which the absorbance maximum is shifted from 415 nm to 410 nm [27]. Non-Mg chelated compounds are not reduced. Note that the shift in retention time observed for $P^{IX}ME$ in the enzyme and control reaction is the results from an aberrance in column running conditions as both compounds retain the absorbance maximum of the $P^{IX}ME$ substrate. (B) Purified *CT*BciA reduces the C8-vinyl group on MgP^{IX} and $MgP^{IX}ME$, as indicated by a complete shift in compound peak absorbance maxima from 415 nm to 410 nm. No activity and correspondingly, no shift in absorbance maximum is observed against non-Mg chelated compounds P^{IX} and $P^{IX}ME$. For abbreviations of substrate names see Table 1.

Similar results were obtained for *CT*BciA, which also reduced only the Mg-chelated porphyrins at a similar rate to that observed for DVP (Figure 17). Also, full conversion of Mg-chelated porphyrins was evident. These results indicate that both 8-vinyl reductases are substrate promiscuous, however activity is dependent upon the presence of the chelated magnesium. This is in contrast to data obtained from *in vivo* experiments, where *CT*BciA was apparently active against P^{IX} and $P^{IX}ME$ (Table 4). It could be that the presence of mvP^{IX} and $mvP^{IX}ME$ in our *in vivo* reactions is actually caused by loss of the magnesium ion during extraction of cell constituents, and that *CT*BciA is not truly active against the non-magnesium chelated porphyrins. Our *in vitro* data clearly show that *CT*BciA is not active against P^{IX} or $P^{IX}ME$, which supports this theory.

The first committed step of BChl biosynthesis is the magnesium chelation of P^{IX} by the magnesium chelatase complex. Metal insertion is believed to tag porphyrin molecules for specific biosynthetic routes leading to e.g. hemes, BChls and corrins. BchSID gene disruptions in *R. capsulatus* result in accumulation of P^{IX}, indicating that the magnesium ion is essential for the correct function of the downstream BChl enzymes (53), and studies in barley show that the presence of a metal ion is essential for reduction of protochlorophyllide (54). It is believed that the magnesium ion is required for the correct orientation and binding of the porphyrin ring in the catalytic pocket of BChl pathway enzymes (55). We therefore suspected that the difference in reaction rates between *RS*BciA and *CT*BciA may be related to different binding or interactions of the substrate(s) with the enzymes.

Circular dichroism analysis of substrate binding

To understand the differences in rates of reaction between *RS*BciA and *CT*BciA, we sought to confirm whether the binding of substrate(s) was similar in the case of both 8-vinyl reductases using CD. Porphyrins and metallo-porphyrins are sensitive CD chromophores, ideal for detecting differences in substrate binding between the two proteins (56). Wavescans in the far UV spectral region showed a change in the structure of *CT*BciA upon addition of MgP^{IX}, with a shift in the characteristic double minima at 208 and 222 nm, associated with α-helical secondary structure (56) (Figure 18A). This change was not observed upon addition of P^{IX} or ZnP^{IX}, further confirming the dependence upon the chelated Mg^{2+} for correct binding of the substrate. Notably, there was no change in the structure of *RS*BciA upon addition of P^{IX}, MgP^{IX}, or ZnP^{IX} (Figure 18B). However, changes in the Soret band of the MgP^{IX} chromophore were apparent upon addition of both *CT*BciA and *RS*BciA.

Studies into the binding of porphyrins by free amino acids indicate that the wavelength of the characteristic peak and inflection switches upon binding by D- or L-isomers of the same amino acid (57). In our case, it could be that distinct side-chain conformational changes are taking place upon binding of MgP^{IX} in *CT*BciA and *RS*BciA, resulting in different changes in the Soret region (Figure 18C, Fig 18D) (57). These data indicate that the two 8-vinyl reductases use a different substrate binding mode, which could be related to their differences in activity.

Interactions of Mg porphyrins with histidine, proline, serine, threonine and tryptophan result in a split CD signal (57). Whether any of these residues are involved in

hydrogen bonding or π stacking of the Mg-chelated porphyrins is not known in the case of 8-vinyl reductase, although it has been suggested that histidine plays a role in coordinating the chelated Mg^{2+} in other BChl pathway enzymes (55). Which residue(s) is involved in coordinating the substrate is not immediately apparent by sequence analyses of *RS*BciA and *CT*BciA (Figure 11), compounded by the fact that no model of the catalytic domain can be obtained. Attempts to crystallize 8-vinyl reductase for 3D structure solution are currently underway, however diffraction quality crystals have not yet been obtained. Analysis of a crystal structure in the presence and absence of substrate could provide detailed information about the substrate binding mode and catalytic mechanism used by 8-vinyl reductase.

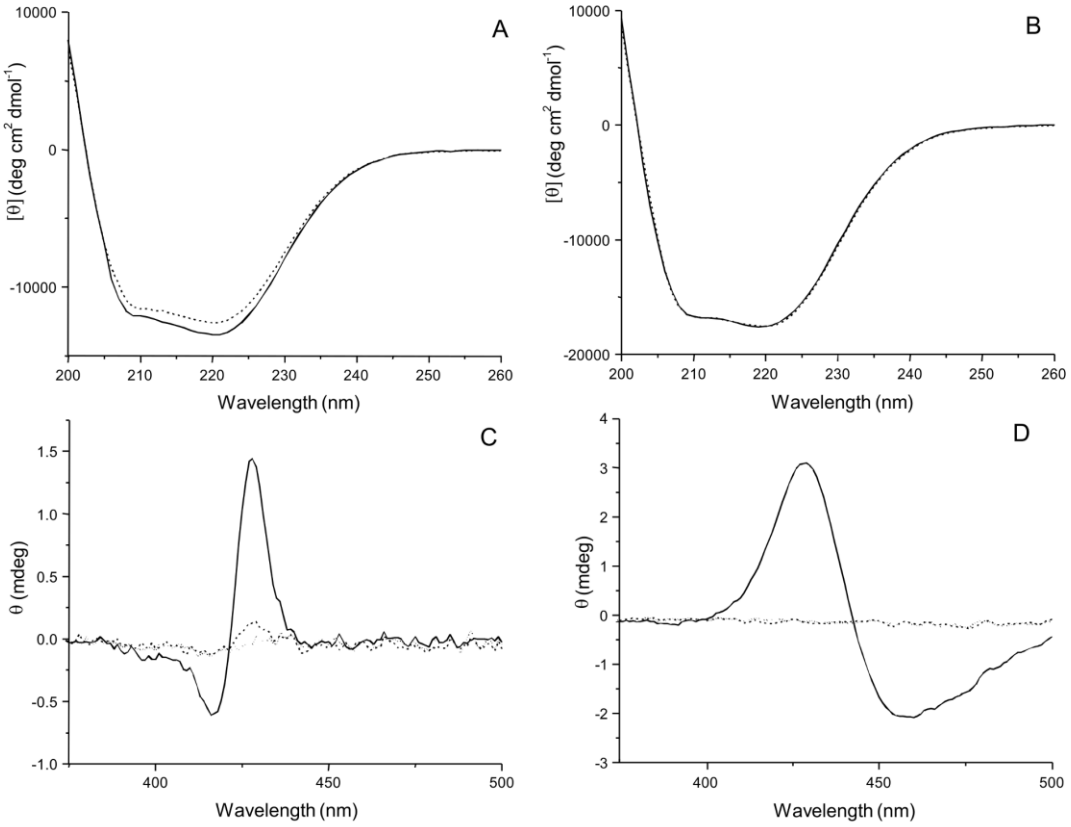

Figure 18.

8-vinyl reductases display structural differences upon binding of MgP^{IX}. CD spectra of purified protein in the far UV region show that **(A)** *CT*BciA and **(B)** *RS*BciA display the characteristic double minima at 222 nm and 208 nm associated with α-helical content (56) (solid line). Upon addition of MgP^{IX} to the protein, a shift is observed in the CD spectrum of *CT*BciA, but not *RS*BciA (dotted line). **(C)** and **(D)** Analysis in the Soret region of MgP^{IX} (dotted line) shows no spectra. Upon addition of purified protein **(C)** *CT*BciA and **(D)** *RS*BciA a change is observed in the Soret band of the porphyrin ring (solid line). The differences in peak and inflection wavelengths may represent MgP^{IX} interactions with different amino acid isomers in the two different proteins (57).

DISCUSSION

8-vinyl reductases are widely distributed in photosynthetic organisms. Putative 8-vinyl reductases have been identified in several different kingdoms of life (Figure 10), where they are suspected to catalyze a key functionalization of porphyrin molecules in (B)Chl biosynthesis (58). Homologous enzymes exist across different species (36), as well as within individual species (59), highlighting the evolutionary selective advantage

afforded by the essential catalytic role of 8-vinyl reductases. Recent studies revealed that homologous plant 8-vinyl reductases are substrate promiscuous, capable of reducing the C8-vinyl group of different pathway intermediates (36). Data in this report demonstrates that substrate promiscuity is not only limited to plant 8-vinyl reductases, but is also a characteristic of 8-vinyl reductases from photosynthetic bacteria. Here, 8-vinyl reductases did not show a particular preference for any pathway intermediate, but activity was entirely dependent upon presence of the chelated Mg^{2+}. These data are in agreement with previous studies (53,54).

It is not fully understood why the presence of a chelated Mg^{2+} in the porphyrin ring is essential for 8-vinyl reductase activity. It is likely that it the metal ion is required for the correct orientation of the porphyrin molecule in the active site or to sterically align the molecule in close proximity with the essential NADPH cofactor (55). Our data provides some insights into the binding of the Mg-porphyrin substrate, which may affect molecular recognition and therefore catalytic activity. These data provide the first evidence that the mode of binding of substrate(s) varies between homologous 8-vinyl reductases. Whether this difference in binding confers a selective advantage to the host, or whether it provides the enzyme with a means to regulate reaction efficiency with differing substrates (36), and therefore pathway flux, remains open to question.

No data exists to show which residues are involved in binding the substrate or those which are involved in catalysis, which could explain the different modes of binding. Prior to this work, 8-vinyl reductase had not been purified, and limited biochemical data had been published (27), therefore the mechanistic details of the

enzyme remained elusive. While this study has shed some light on the reduction of the C8-vinyl group of BChl intermediates by different 8-vinyl reductases, an in-depth analysis of the step-by-step catalytic process is still required. Detailed comparative biochemical and structural characterization of homologous 8-vinyl reductases would provide the information needed for a full understanding of the reaction mechanism and the substrate recognition of this enzyme.

The purpose of this study was to extend the engineered BChl biosynthetic pathway in *E. coli* (5). It was necessary to insert a downstream enzyme that could accept the multiple products of our existing Mg chelatase-methyl transferase system (15). In our endeavours to select a suitable candidate for this role, we discovered that *CT*BciA is capable of reducing the C8-vinyl group of several substrates when expressed as part of our engineered pathway. This is a significant step toward our goal of recreating the full BChl pathway in *E. coli*. One of the challenges of building a pathway engineered microbe is maintaining balance in pathway flux. Often, a slow catalytic rate of one or more enzymes can result in pathway bottlenecks, or unwanted side reactions can lead to inefficient use of metabolically expensive molecules (60). We have not yet eliminated the potential for the production of non Mg-chelated "dead end" porphyrins, although it is likely that the presence of mvP^{IX}ME actually results from loss of the Mg ion from mvMgP^{IX}ME upon extraction from the cell. Nonetheless, the addition of a substrate promiscuous 8-vinyl reductase to our system does provide a shuttle for the Mg-chelated intermediates, resulting in a reduction in the preference to produce P^{IX}ME, and altering

the pathway balance to produce equal ratios of the mono-vinyl Mg-chelated porphyrins (Table 4).

We were surprised to find that the closely related homologue *RS*BciA was inactive in our pathway. Furthermore, the *in vitro* catalytic rate of *RS*BciA was too slow to be relevant for our engineering purposes, despite extensive attempts to optimize reaction conditions. It may be that *RS*BciA requires an unknown species-specific cofactor, chaperone, interacting partner or shuttling enzyme for efficient function, which is the case for other enzymes involved in BChl biosynthesis (18,61,62). It has been suggested by others that BchJ is not a 8-vinyl reductase as previously indicated (49), but that it functions as a carrier or shuttle for porphyrin intermediates (18). However, we found that coexpression of *RS*BchJ had no effect on *RS*BciA activity. Future studies beyond gene-knockouts could elucidate the exact nature of BciA behavior in *R. sphaeroides*, and could clarify whether this particular 8-vinyl reductase is capable of acting alone or whether its activity is upregulated in the presence of certain other members of the pathway or under different reaction conditions (16). Very recently, an anaerobic 8-vinyl reductase (BciB) from the green sulfur bacterium *Chloroherpeton thalassium* was characterized. BciB requires two (4Fe-4S) clusters, FAD, and a reductant such as ferredoxin or sodium dithionite to reduce the C-8 vinyl group of DVP (63). This study reveals the mechanistic diversity of 8-vinyl reductases, and highlights the importance of *in vitro* characterization for a full appreciation of optimal conditions for catalysis. Furthermore, the presence of a not-yet-identified 8-vinyl reductase in *R.*

sphaeroides cannot be ruled out, as we are only just gaining insights into the sequence diversity of this class of enzymes (59).

Relying upon sequence analyses and the relatively limited biochemical data that existed for 8-vinyl reductases was not sufficient for the strategic design of our engineered BChl system in *E. coli*. We were not able to predict that *CT*BciA would be active in our engineered pathway and that *RS*BciA would be inactive in the same engineered pathway. The unsuitability of *RS*BciA for our purposes only became truly apparent upon our own detailed biochemical and biophysical characterization of the two enzymes. In some ways, this study serves as a good example to underline the fact that the strategic and streamlined design and engineering of metabolic pathways is heavily dependent upon having a detailed knowledge of the catalytic mechanism and/or three dimensional structure of the enzyme(s) in question (64). Sometimes this data is not available to the pathway engineer, in which case it becomes necessary to characterize the enzymes involved, and gather the detailed information needed to optimize the system for a particular purpose. Pathway engineering is not a straightforward process of building a chain of enzymes to make a product, rather it is the intricate design, creation and polishing of a living system to entice it to carry out a completely new activity.

In conclusion, this study provides an in-depth characterization of two 8-vinyl reductases from two photosynthetic organisms, and gives insights into the potential diversity of function with regards to substrate promiscuity and binding of substrates. This study brings us a step closer to the realization of the creation of an industrially

relevant synthetic microbe that can use sunlight as a cheap source of energy to drive the biosynthesis of valuable and designer target molecules.

Chapter 4. A first look at the Mg protoporphyrin IX methylester cyclases from *Rhodobacter sphaeroides*.

This chapter is adopted from: A first look at the Mg protoporphyrin IX methylester cyclases from *Rhodobacter sphaeroides*. Ilya B. Tikh, Maureen B. Quin and Claudia Schmidt-Dannert. In preparation for publication in Applied Environmental Microbiology.

Chapter 4 Overview

Porphyrin and functionalized porphyrin derivates are sought after for their use in artificial solar cells, cancer therapies and biosensor applications. We have previously demonstrated the overexpression of the heme biosynthetic pathway in *E. coli* yielding high levels of porphyrin molecules, including protoporphyrin IX (PIX) and heme. Some of the aforementioned applications would benefit from more diversely functionalized set of porphyrin backbones. To achieve this, we are in the process of combining the heme and the bacteriochlorophyll biosynthetic pathways in a heterologous host for high level expression. We previously demonstrated functional expression of magnesium chelatese and methyl transferase, the first two enzymatic steps in bacteriochlorophyll biosynthesis, in *E. coli.* The next step in the pathway is the conversion of the Magnesium protopophyrin IX monomethylester (MgPIX ME) to divinyl-protochlorophyllide, by cyclization of the characteristic fifth ring. However, none of the cyclase enzymes thought to perform this reaction have ever been purified and characterized. We therefore set out to characterize the aerobic cyclase AcsF and anaerobic cyclase BchE, from *R. sphaeroides*. We found that AcsF did not express in *E. coli* under a constitutive promoter, and that BchE was expressed in an insoluble fashion. Nonetheless, we demonstrate the first purification of BchE from the native host *R. sphaeroides,* and confirm by EPR the presence of a 4Fe-4S cluster in BchE, as well as interactions with intermediates of

bacteriochlorophyll biosynthesis and SAM, the predicted co-factor of the cyclase. We also demonstrate the successful of BchE purification from the heterologous host *E. coli*, which is an important tool for pathway engineering. Ultimately, these results bring us significantly closer to understanding the conversion of MgPIX ME to DVP and our goal of reconstituting bacteriochlorophyll biosynthesis in a heterologous host, thus creating a system for the production of a diverse set of highly functionalized porphyrin molecules.

INTRODUCTION

Photosynthetic organisms compose a large and vital portion of the earth's ecosystem. As a result, photosynthesis, specifically the mechanism of light capture, has been extensively studied. The process of light capture is known to rely on a core set of proteins composing the cellular light harvesting machinery, complexed with accessory proteins and pigments (reviewed in (1)). The pigments involved in photosynthesis are (Bacterio)Chlorophylls ((B)Chls) and carotenoids (1, 2). (B)Chls are essential for efficient light capture, and carotenoids play multiple roles, including photoprotection (3, 4). In addition to being necessary for photosynthesis, tetrapyrrole based molecules, including modified porphyrin intermediates of BChl biosynthesis, have interesting industrial applications such as light harvesting molecules in artificial photosynthetic systems and potential cancer therapies (5-8).

Given the relative importance of BChls in efficient photosynthetic processes, much research has focused on the BChl biosynthetic pathway. The enzymes responsible for most of the catalytic steps have been identified, purified and their function has been

confirmed using standard biochemical methods. The only exception is the enzyme responsible for the conversion of Mg protoporphyrin IX monomethylester (MgPIX ME) into divinyl protochlorophyllide (DVP) (Figure 19) (9). The characteristic isocyclic fifth ring of DVP is the most pronounced feature that distinguishes BChls from other tetrapyrrole molecules. Genetic studies have identified two putative enzymes catalyzing the cyclization of this ring on MgPIX ME. One of these cyclases is functional under aerobic conditions (AcsF), and the other is functional under anaerobic conditions (BchE) (10-12).

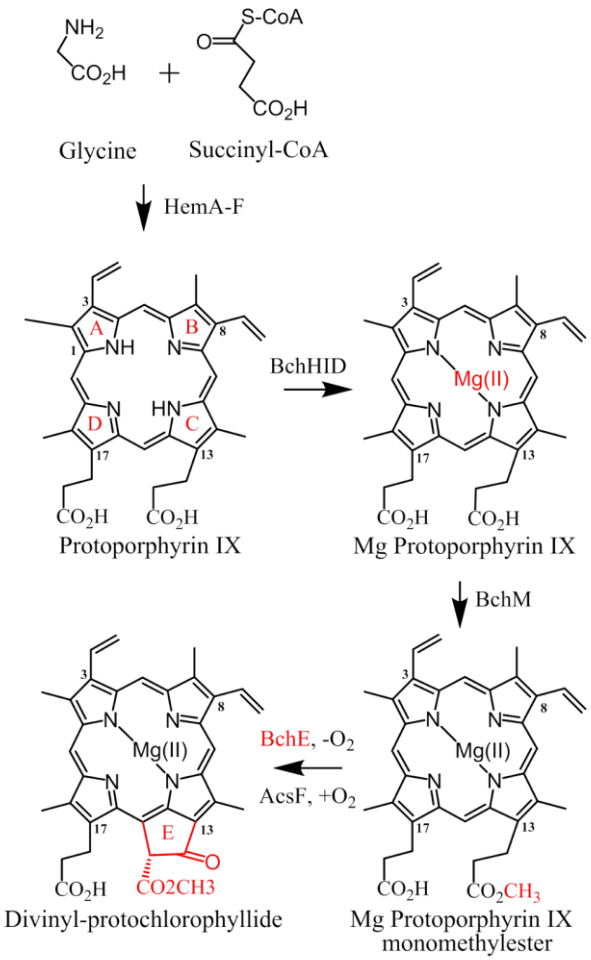

Figure 19.

Production of Mg Protoporphyrin IX Methylester. Bacteriochlorophyll biosynthesis shares a set of common precursors with the heme biosynthetic pathway. Specifically, HemA-F enzymes convert glycine and succinyl-CoA into protoporphyrin IX (PIX) (23). A magnesium ion is then chelated into Protororphyin IX by BchHID complex in the first commited step of bactiochlorophyll biosynthesis, producing Mg-protoporphyrin IX (MgPIX) (22). Subsequently, MgPIX is converted into the Mg protoporphyrin IX monomethylester (MgPIX ME) in a SAM dependent fashion by BchM (22). The conversion of MgPIX ME into DVP is presently thought to be performed AcsF under aerobic conditions or BchE under anaerobic conditions (10).

The putative aerobic cyclase, AcsF, is found in plants, algae and oxygen tolerant photosynthetic bacteria (9, 11, 13). Several AcsF homologues have been identified as one of a number of membrane associated components required for the aerobic conversion of MgPIX ME to DVP (11, 14, 15). In addition to AcsF and potentially a second membrane, a soluble unidentified accessory protein is required for aerobic cyclase activity (11, 16). While multiple AcsF homologues have been identified, only a small amount of information is available about the enzyme activity. Based upon sequence analysis, AcsF is predicted to contain a ferritin like domain (17). Furthermore, AcsF contains two DExxH motifs characteristic of diiron carboxylate proteins (Figure 20) (10, 18). Lastly, it has been determined that the oxygen incorporated during the formation of the isocyclic ring under aerobic conditions, and presumably by AcsF, comes from a source other than water (19).

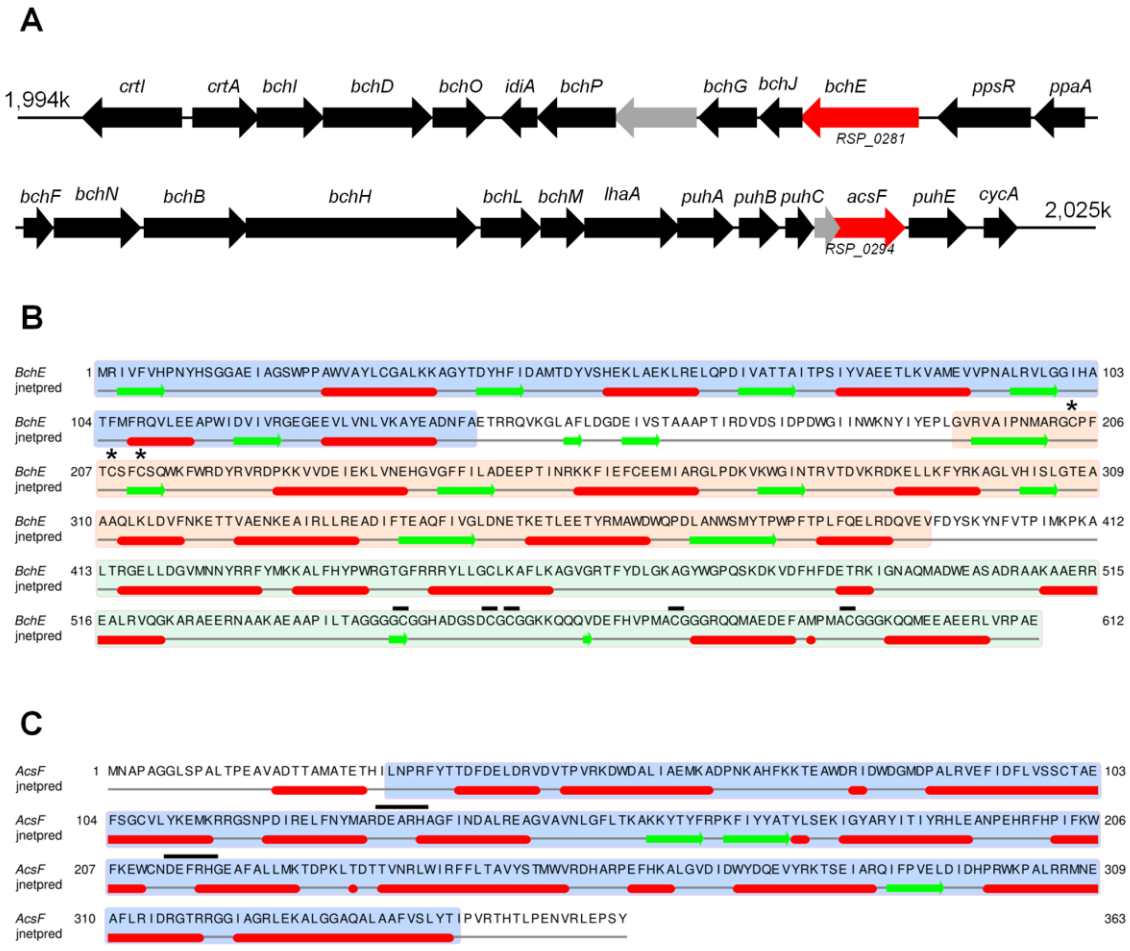

Figure 20.

Overview of the *R. sphaeroides* photosynthetic gene cluster and two putative cyclases. (A) A scale representation of the photosynthetic gene cluster from *R. sphaeroides* with the two putative cyclase genes represented in red. Genes with hypothetical function are shown in gray and the numbers at each side of the operon represent chromosomal position in thousands of base pairs. **(B)** The protein sequence of BchE was analyzed for any known domains as well as for secondary structure. The sequence highlighted in blue represents the predicted cobalamin binding domain ($BchE_{1-142}$). Sequence highlighted in beige corresponds to predicted 4Fe-4S domain ($BchE_{192-393}$) with the conserves cysteines forming the CXXXCXXC 4Fe-4S binding motif marked with asterisks. The sequence highlighted in green is the C-terminus domain ($BchE_{413-612}$) with cycteins in the region of low predicted secondary structure marked by black bars. Green arrows represent predicted α-helices, while red bars represent predicted β-sheets as analyzed by Jpred (28). **(C)** Similarly, the sequence of AcsF was analyzed for known domains and secondary structure. A ferritin-like domain, highlighted in blue, covers the majority of AcsF and two conserved DEXXH motifs present in diiron oxygenases are noted by black bars (18). The highly helical secondary structure is also common among diiron oxygenases (18). Domains were predicted using the NCBI Conserved Domain Database (17).

The anaerobic cyclase, BchE, is found in bacteria able to perform photosynthesis under anaerobic conditions, such as the model *Rhodobacter* species. Little is known about BchE, although it is remarkably well conserved across purple and green sulfur bacteria (Figure 21). BchE was first identified using a transposon screen from mutants of *R. capsulatus* accumulating large quantities of MgPIX ME (12, 20). The majority of the information we have on BchE, such as it likely being a 4Fe-4S radical SAM enzyme and containing a cobalamin binding motif, has been derived from sequence analysis. Its activity has not yet been fully confirmed in vitro. The importance of cobalamin for BchE activity was experimentally verified *in vivo* using a cobalamin deficient *R. capsulatus* strain and feeding experiments (21). Furthermore, isotopically labeled water appears to be the source of oxygen incorporated into DVP by purple bacteria grown under anaerobic conditions, suggesting that the BchE reaction mechanism is distinct from that employed by AcsF (19). A single report has been published regarding an attempt to express *bchE* in *E. coli*, resulting in inclusion bodies, thus functional purification has not yet been demonstrated. Therefore, the biochemical details of BchE remain unknown (10).

We successfully rebuilt both the heme and the bacteriochlorophyll pathway, resulting in full biosynthetic production of MgPIX ME, in *E. coli*. This was achieved using a modular approach combining genes from a diverse range of sources (22, 23). The next critical step in our pathway engineering is the formation of the isocyclic fifth ring by the as-yet-uncharacterized cyclase. As both of the putative cyclases, AcsF and BchE, potentially require chaperones and numerous cofactors, we initially employed our *in vivo* system. We hypothesized that by including either of the putative cyclases to our already

reconstructed BChl pathway, *E. coli* would produce DVP. The results here show that neither enzyme is active in our pathway engineered *E. coli*. AcsF proved to be difficult to express at detectable levels in our recombinant system, while BchE was expressed solely in inclusion bodies. In order to understand the mechanism of the cyclization reaction, and therefore optimize our pathway, we set out to purify the anaerobic cyclase BchE, and to characterize its activity *in vitro*.

This work represents the first published study that explores the suitability of both AcsF and BchE in an *in vivo* reconstitution of the BChl biosynthetic pathway to yield DVP. Furthermore, for the first time, we present the biophysical characteristics of purified BchE from *R. sphaeroides* as well as reconstituted BchE purified from *E. coli*. We are able to confirm the presence of the 4Fe-4S cluster in BchE and its interaction with SAM as a cofactor. While catalytic turnover was not detected, we demonstrate that there is an interaction between BchE and MgPIX ME. Taken together, this data supports previously published genetic evidence that BchE plays a role in the conversion of MgPIX ME to DVP (20). This information significantly furthers our understanding of the cyclization of the characteristic fifth ring of DVP. Finally, because of the new information presented here, we are now one step closer to understanding the crux reaction of BChl biosynthesis that has thwarted scientists for decades.

Figure 21.

Sequence alignments of BchE homolgues. Alignment of several close BchE homologues from closely related α-proteobacteria as well as BchE homologues from a β-proteobacterium and a green sulfur bacterium used to illustrate high homology of BchE between different organisms. Conserved residues highlighted in blue. The region with highest variability between the homologues is found at the C-terminus. Interestingly, only BchE from *R. sphaeroides* contains a cystein and glycine rich region inserted near the C-terminus of the enzyme (highlight with a red box).

MATERIALS AND METHODS

Chemicals and enzymes

The chemicals used in this study were purchased from Sigma-Aldrich (St. Louis, MO), unless otherwise noted. Restriction enzymes and polymerases were purchased from NEB and used in accordance with manufacturer instructions. *R. sphaeroides* ΔRCLH was a kind gift from Dr. Beatty (24)

Gene Cloning

Full length *bchE* (NCBI Gene ID: 3719193) was amplified from *R. sphaeroides* 2.4.1 genomic DNA using primers P1 and P2. The PCR product and pUCBB-ntH6-GFP (25) were both digested with *Nde*I and *Not*I, gel purified and ligated to create pUCBB-ntH6BchE$_{1\text{-}162}$, which was sequenced to confirm a lack of mutations. BchE$_{192\text{-}393}$ was created by amplifying the region containing amino acids 192-393 from pUCBB-ntH6BchE with primers P3 and P4, with the 6xHis N-Terminus tag built into primer P3. The PCR fragment and pET30b were digested with *Nde*I and *Not*I, gel purified and ligated together to create pET30-ntH6BchE$_{192\text{-}393}$. Similarly, plasmid pET30-BchE$_{413\text{-}612}$ was created by amplifying the region containing amino acids 413-612 with primers P5 and P6. Primer P6 contained an inclusion body localization tag (26). The PCR product was subsequently ligated into the *Nde*I and *Not*I sites of pET30. Plasmid pBBRBB-*Ppuf*$_{843\text{-}1200}$-BchE-ctH6 was created by amplifying *bchE* with primers P7 and P8,

digesting with *Bgl*II and *Not*I and ligating into the same site of pBBRBB-*Ppuf*$_{843\text{-}1200}$-DsRed (Tikh *et. al* 2013, submitted). BchE was cloned into the *Bgl*II and *Not*I sites of pUCBB-GFP by amplifying with primers P9 and P10, digesting both the PCR product and the vector with the same enzymes, and ligating the two DNA fragments to form pUCBB-*bchE*. Vector pUCBB-*acsF* was constructed in a similar manner with *acsF* amplified with primers P11 and P12 and cloned into the *Nde*I and *Not*I sites. Gene stacking to create pCDFBB-*bchM-bchE* and pCDFBB-*bchM-acsF* was perfomed following standard BioBrick™ protocols. In addition, N and C terminus 6xHis-tags were added to the proteins by subcloning into other BioBrick™ vectors. Primers used in this study are shown in Table 5. Bacterial cultures were grown in LB medium, unless otherwise noted, and supplemented with kanamycin (30 µg/µl) and ampicillin (100 µg/µl) as needed for plasmid maintenance.

In vivo production of porphyrins

Production of magnesium porphyrins, including MgPIX ME, in *E. coli* cells was achieved utilizing plasmids pUCMOD-*SID*, pAC-*hemAD*, pBBRBB-*hemEF* and pCDFBB-*bchM* as previously described (22, 23)(Tikh et al 2013, submitted). Extraction and HPLC analysis of porphyrin derivatives from cells was accomplished using 7:2:1 acetone:methanol:water and analyzed on a Zorbax 300SB C-18 reverse phase column (Agilent) using the Agilent 1100 series HPLC. The running conditions were 80%-100% of methanol over 20 minutes with the aqueous phase composed of 0.1M ammonium acetate pH 5.1. To determine *in vivo* activity of BchE and AcsF, plasmid pCDFBB-*bchM* was replaced with one of the following: pCDFBB-*bchM-bchE* or pCDFBB-*bchM-acsF*.

BchE sequence analysis

Sequence alignments of BchE homologues were performed in MEGA version 5 using the ClustalW algorithm with default settings (27). Secondary structure predictions for BchE and AcsF were obtained using the automated prediction mode from the Jpred 3 server (28). Conserved domains were identified using the NCBI Conserved Domain Database (17). Operon structure for *R. sphaeroides* 2.4.1 BChl biosynthetic genes was deduced from NCBI gene annotations.

Expression and purification of BchE in Rhodobacter

Plasmid pBBRBB-*Ppuf*$_{843\text{-}1200}$-BchE-ctH6 was introduced into *R. sphaeroides* ΔRCLH (24) strain, containing a deletion of all components of the reaction center, via conjugation, as previously described (29). For protein expression, a Manifors 3L bioreactor containing Sistrom's minimal medium (30) supplemented with kanamycin was made anaerobic by flushing with N_2 gas and inoculated with 200 ml of culture in exponential phase. The bioreactor was maintained at 30°C with stirring at 200 rpm throughout the experiment. Succinate (30 mM) and DMSO (50 mM) were used as the electron donor and electron acceptor, respectively. The cells were harvested by centrifugation after 18 hours of growth and were stored at -80°C until use. All subsequent steps were performed under anaerobic conditions inside a Coy anaerobic chamber or under an argon stream, unless otherwise noted. Cells were resuspended in anaerobic lysis buffer (50 mM Tris-HCl pH 8, 100 mM NaCl, 5 mM imidazole) and lysed by sonication. After lysis, SDS was added to a final concentration of 0.1% and allowed to incubate for 1 hour at 4°C, with gentle stirring. Following the incubation, cell lysate was centrifuged in

a sealed centrifuge bottle at 19,000 g to remove cell debris. Clarified cell lysate was brought back into the anaerobic chamber and loaded onto a 1 ml HisTrap FF crude column (GE Lifesciences) equilibrated with buffer A (50 mM Tris-HCl pH 8, 100 mM NaCl, 0.1% Triton X-100, 5 mM imidazole). The column was subsequently washed with 5 ml of buffer A containing 5, 50 and 100 mM imidazole. The protein was eluted with 3 ml of buffer A containing 250 mM imidazole. Purified protein was analyzed by SDS-PAGE. UV/Vis spectra of the protein samples was taken in sealed cuvettes on an Agilent 5483 spectrophotometer, anaerobic dithionite was titrated into the protein solution using a gas tight syringe at 100 µM increments.

Expression, purification and refolding of BchE in E. coli

E. coli BL21 (DE3) were used to express both the full length $BchE_{1-612}$ as well as $BchE_{192-393}$ and $BchE_{413-612}$. The protein was expressed as inclusion bodies, which were separated from *E. coli* cell debris and were resuspended in solubilization buffer (6 M urea, 50 mM Tris pH 8) as previously described (22). Urea solubilzed proteins were refolded by drop wise addition into a 10 fold excess of buffer B (50 mM Tris-HCl pH 8, 20% glycerol) in the presence of 0.1 M DTT, 200 µM $FeCl_3$, 200 µM $Fe(NH_4)(SO_4)_2$ and 200 µM Na_2S under strictly anaerobic conditions and allowed to gently stir for 20 min on ice (31). Following the incubation, refolded proteins were centrifuged for 20 minutes at 19,000 g and 4°C. Soluble refolded proteins were partially purified by loading onto G25 coarse resin (GE Healthcare) and eluting with buffer C (50 mM Tris-HCl pH 8, 10% glycerol). Fractions containing the characteristic brown color were checked by UV/VIS

and SDS-PAGE, from those, fractions containing >80% of BchE were pooled, concentrated with an Amicon spin unit (Millipore) and prepared for EPR.

EPR analysis

X-band EPR spectra were collected with a Bruker Elexsys E-500 spectrometer with a Bruker dual mode cavity and Oxford Instruments ESR 910 cryostat. Cryogenic temperature control was achieved with an Oxford Instruments ITC-503s temperature controller. EPR spectra collected with the following conditions: temperature 20K, power 200µW, gain 60, modulation amplitude 10 G, modulation frequency 100 G.

BchE activity assay

MgPIX, MgPIX ME and PIX ME were extracted from *E. coli* BL21(DE3) cells containing pUCMOD-DIS and pCDFBB-BchM using 1:7:2 of water:acetone:methanol as previously described (22) (cite reductase), dried down and stored at -20°C until futher use. Prior to use, the pigment mixture was resuspended in assay buffer (50 mM Tris-HCl pH 8, 250 mM NaCl, 0.3% Triton X-100). The final assays contained a mixture of MgPIX, MgPIX ME, PIX ME, 500 µM SAM, 500 µM NADH, 0.3% Triton X-100. The activity assays were made anaerobic by flushing each component with argon gas and maintained under strict anaerobic conditions in a Coy anaerobic chamber. The final enzyme assays, performed under anaerobic conditions, contained 80uL with various concentrations of purified BchE or buffer and allowed to incubate at room temperature for 24 hours, at which point they were taken out of the anaerobic chamber and quenched with 400uL of ice cold acetone. The quenched reactions were centrifuged at 21,000 g and run through a Zorbax 300SB C-18 reverse phase column (Agilent) using the Agilent 1100

series HPLC. The running conditions were 80%-100% of methanol over 20 minutes with the aqueous phase composed of 0.1M ammonium acetate pH 5.1. Reaction products were monitored at 400, 412, 420 and 440 nm.

Table 5. Primers used in Chapter 4.

P1	5′- GCCATATGATGCGTATCGTATTCGTTCACC-3′
P2	5′-TAGCGGCCGCTCACTCCGCCGGACGGACGAG-3′
P3	5′-ATCATATGCACCATCACCATCACCATGGCGTGCGCGTCGCGATCC-3′
P4	5′-TAGCGGCCGTCACGACCTGATCGCGCAGCTCCTGG-3′
P5	5′-ATCATATCTGACCCGCGGCGAGTTGCTCG-3′
P6	5′-AGCGGCCGCTCAATCTTTCAGCAGCAGCAGCAGCAGGGTCGGCTCCGCCGGACGGACGAG-3′
P7	5′-TATAGATCTATGCGTATCGTATTCGTTCACC-3′
P8	5′-TAGCGGCGGCTCAGTGATGGTGATGATGCTCCGCCGGACGGACGAG-3′
P9	5′-TATAGATCGATGCGTATCGTATTCGTTCACC-3′
P10	5′-TAGCGGCCGCTCACTCCGCCGGACGGACGAG-3′
P11	5′-GCCATATGATGAACGCGCCGGCCGG-3′
P12	5′-AGGCGGCCGCTCAATAGCTCGGCTCCAGTCGGACG-3′

RESULTS

In vivo activity of BchE and AcsF as part of the BChl pathway

Given our previous success reconstructing the heme and the BChl biosynthetic pathways using genes from the model purple bacterium *R. sphaeroides* as well as other sources (22, 23), and the ease of gene stacking afforded by the BioBrick™ system, we initially introduced both of the putative MgPIX ME cyclase genes from *R. sphaeroides* directly into the pathway. Either BchE or AcsF was expressed in *E. coli* cells already expressing HemA-F as well as BchSID and BchM, and the porphyrins produced by this system *in vivo* were monitored. Cultures expressing AcsF were grown under aerobic conditions, as AcsF is predicted to encode the aerobic MgPIX ME cyclase (10, 11). Given that BchE is predicted to encode an anaerobically active enzyme, cultures expressing BchE were grown under both aerobic and anaerobic conditions. Ultimately,

none of the cultures exhibited any conversion of MgPIX ME into DVP (Figure 22, Figure 23). Given the robustness we previously observed with our porphyrin expression system, we were surprised that neither of the enzymes exhibited cyclase activity. AcsF is known to require an undefined soluble component for activity (11), which we were not able to identify from the genomic context of AcsF in *R. sphaeroides* (Figure 20). However, prior to making any conclusions regarding potentially missing maturation machinery or cofactors, we decided to verify expression levels of each putative cyclase in *E. coli* cells.

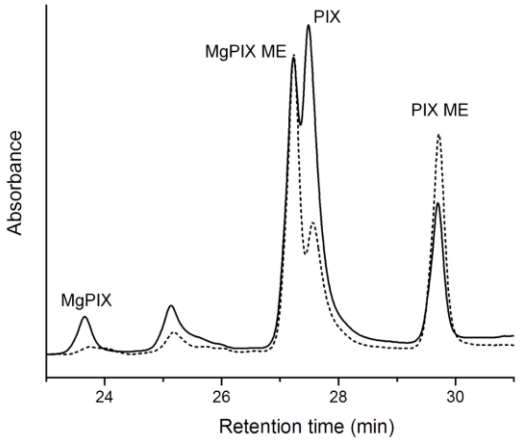

Figure 22.

Reconstitution of the *in vivo* BChl biosynthetic pathway with the addition of AcsF. *E. coli* cells were grown with plasmids pUCMOD-*SID* and pCDFBB-*bchM* in the control sample (solid trace) in order to overproduce magnesium porphyrins, including MgPIX ME, the predicted substrate for AcsF. Alternatively, cultures were grown containing vectors pUCMOD-*SID* and pCDFBB-*bchM-acsF* (dashed trace). Pigments from both sets of cultures were extracted, analyzed by HPLC and detected at 420 nm. The unlabeled peak is known to belong to zinc porphyrin and is sometime present during *in vivo* experiments. While some variation in the amount of different porphyrins produced in observed, no peak corresponding to DVP is present in either of the two cultures. For the location of DVP, see figure 23.

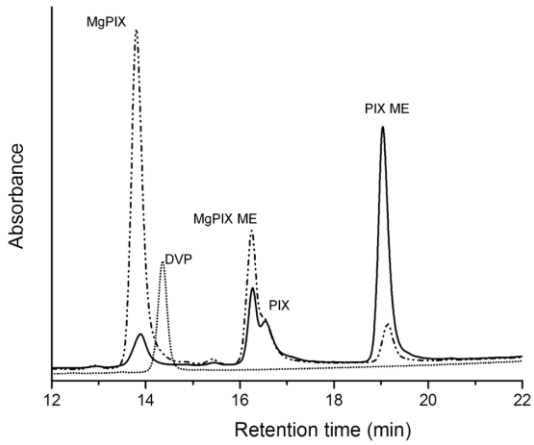

Figure 23.

Reconstitution of the *in vivo* BChl biosynthetic pathway with the addition of BchE. *E. coli* cells were grown anaerobically with plasmids pUCMOD-*SID* and pCDFBB-*bchM* in the control sample (solid trace) in order to overproduce magnesium porphyrins, including MgPIX ME, the predicted substrate for BchE. Alternatively, cultures were grown containing vectors pUCMOD-*SID* and pCDFBB-*bchM-bchE* (dashed trace). Dotted trace shows the location of DVP standard. Pigments from both sets of cultures were extracted, analyzed by HPLC and detected at 420 nm. While some variation in the amount of different porphyrins produced in observed, no peak corresponding to DVP is present in either of the two cultures.

Expression of the putative cyclase genes in *E. coli*.

In order to assess cellular levels of AcsF and BchE and to potentially purify the enzymes, a set of constructs were created containing either a C or an N terminus 6xHis-tag, for purification purposes. Lysates from cultures transformed with these constructs were analyzed by Western blotting in order to confirm the expression of the cyclases (Figure 24A). We found that BchE containing an N-terminus tag was well expressed, but was present in the insoluble fraction, confirming previous reports (Figure 24B) (10). Interestingly, BchE containing a C-terminus tag was found to be rapidly degraded within the cells (Figure 24A). Since BchE is predicted to contain an Fe-S cluster, we also attempted to coexpress it with pACYC-*isc*, a chaperone vector containing Fe-S maturation machinery (32). However, the presence of the chaperone vector did not result in any improvement in the amount of soluble BchE produced.

In the case of AcsF, no protein was detected in the *E. coli* host cells (Figure 24A). The lack of soluble protein production for BchE and the lack of detectable AcsF expression levels may explain the lack of activity demonstrated during our *in vivo* assays. Given that we were not able to detect any AcsF in *E. coli* host cells during our expression trials and that we were not able to identify the putative soluble component believed to be required for AcsF activity (11), we focused our efforts on characterizing the well-expressed, albeit insoluble, anaerobic cyclase BchE.

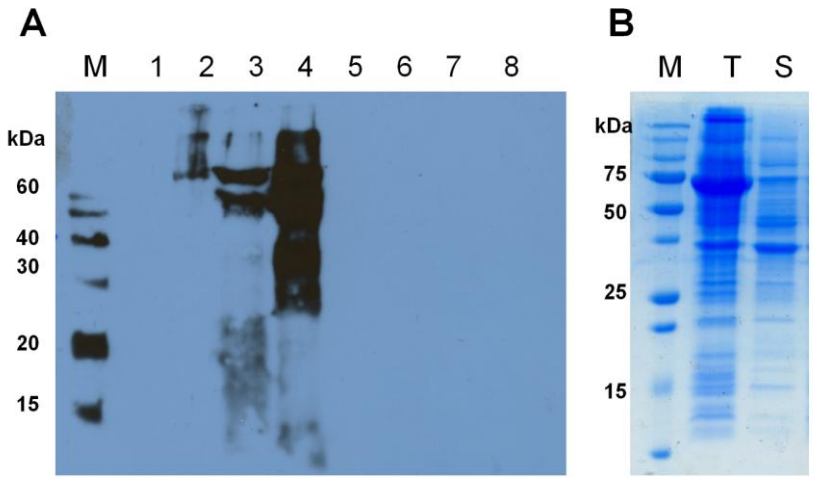

Figure 24.

Analysis of BchE and AcsF levels in *E. coli*. (A) Western blot analysis expressing BchE and AcsF with N and C terminus 6xHis-tags. Odd lanes represent soluble fractions and even lanes are from the insoluble cell lysates. Samples are as follows: molecular weight marker, M; BchE with an N-terminus 6xHis-tag, 1-2; BchE with a C-terminus 6xHis-tag, 3-4; AcsF with an N-terminus 6xHis-tag 5-6; AcsF with a C-terminus 6xHis-tag 7-8. Interestingly, it appears a BchE with a C-terminus tag is rapidly degraded within the cells and while BchE with an N-terminus tag is stable, it is present only in inclusion bodies **(**see gel **B)**. Notably, we did not observe any AcsF under our culture conditions. **(B)** SDS-PAGE of $BchE_{1-612}$ with an N-terminus 6xHis-tag. All of the protein is insoluble and present in inclusion bodies.

BchE sequence analysis.

To better understand BchE prior to attempting any further experiments, detailed sequence analysis was performed. The sequence of BchE revealed a cobalamin binding domain between amino acids 1-142 (henceforth referred to as $BchE_{1-142}$), which correlates well with *in vivo* studies suggesting cobalamin is required for the activity of the cyclase (21). Following the cobalamin binding domain, amino acids 192-393 contain a conserved CXXXCXXC sequence associated with 4Fe-4S containing radical SAM enzymes, which utilize S-adenosyl-L-methionine (SAM) as a cofactor (Figure 20) (33, 34). The region containing the putative 4Fe-4S cluster is predicted to have a well defined secondary structure (Figure 20) and will be referred to as $BchE_{192-393}$ from this point on. In addition to the two well defined domains, $BchE_{1-142}$ and $BchE_{192-393}$, the final 199

amino acids have been termed $BchE_{413-612}$ and contain a glycine rich region, similar to those implicated in nucleotide binding in other enzymes (35-37) and a cysteine rich region ($CX_8CXCX_{17}C$) with some similarity to motifs known to be involved in the formation of Fe-S clusters (Figure 20). This region appears to be unique to R. sphaeroides (Figure 21). While the sequence analysis suggests the presence of a 4Fe-4S cluster within $BchE_{192-393}$, it was imperative to experimentally confirm its presence and composition as maturation requirements can vary between different types of Fe-S clusters. As we were especially interested in the composition of the Fe-S cluster, for subsequent experiments we chose to utilize the full length $BchE_{1-612}$ as well as $BchE_{192-393}$ and $BchE_{413-612}$.

Purification of BchE from *R. sphaeroides*.

Due to the presence of a predicted 4Fe-4S cluster and the complex machinery required for Fe-S cluster maturation, initial purification of BchE was carried out using *R. sphaeroides*, the native host, in order to maximize the yield of fully matured protein. Sequence analysis of BchE did not suggest the presence of any transmembrane regions, so our assumption was that BchE would be soluble when purified from *R. sphaeroides*. Initial purification attempts using soluble cell lysate resulted in the presence of a mostly pure band of approximately 60 kDa upon the elution of the Ni+ column with 250 mM imidazole (Figure 25A). As BchE is predicted to be 70 kDa, the band was excised from the SDS-PAGE, digested with trypsin and analyzed by tandem mass-spectrometry. The mass-spectrometry analysis revealed that the majority of the peptides identified belonged to *R. sphaeroides* GroEL (60 kDa), although several peptides assigned to BchE were also

detected. N-terminus peptide sequencing confirmed GroEL was indeed the 60 kDa protein purified through affinity chromatography (Table 6).

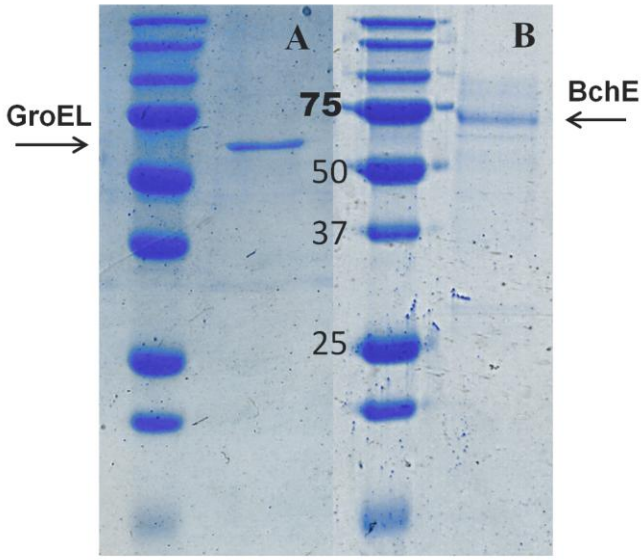

Figure 25.

Purification of BchE from *R. sphaeroides*. SDS-PAGE showing the 250mM imidazole elution from a Ni^{+} column loaded with cell lysate from *R. sphaeroides* expression C-term histidine tagged BchE that was either untreated **(A)** or solublized with a detergent **(B)** prior to loading onto the column. Expected sizes for GroEL and BchE are 60kDa and 70kDa, respectively.

Table 6. N-terminus peptide sequence of the 60 kDa band from Figure 25A.
The N-terminus peptide sequencing was performed at the University of Minnesota Biomedical Genomics Center.

Cycle	Residue	pMole	Comments
1	A	17.523	
2	A	12.104	
3	X	---	X=Cys or modified AA
4	D	5.956	V
5	V	10.886	
6	X	---	X=Cys or modified AA
7	F	8.797	
8	D	5.42	
9	T	5.409	G
10	D	5.762	A

In order to separate BchE from the GroEL complex, detergent was added to the cell lysate (38). Subsequent purification resulted in a single band of approximately 70 kDa on an SDS-PAGE gel, the predicted size of BchE (Figure 25B). The isolated protein was light brown in color, and displayed UV/Vis absorbance features at 350 nm and 415 nm, both characteristic of fully mature 4Fe-4S containing enzymes (Figure 26A) (31, 39). Furthermore, as expected for functional 4Fe-4S clusters, a decrease of both features was evident upon the reduction of the cluster from the 2^+ to the 1^+ state by dithionite (Figure 26A).

Given that some Fe-S enzymes are stable and functional in the presence of oxygen (40, 41) we attempted the same purification protocol under aerobic conditions. Aerobic purification yielded similar quantities of BchE, however, the UV/Vis trace of aerobically purified enzyme was strikingly different and suggested the presence of a porphyrin molecule (Figure 26B). The peak at 420 nm from aerobically isolated BchE is relatively sharper and significantly more intense than any expected 4Fe-4S features. Furthermore, the reduction of the aerobic sample revealed two peaks between 500-600 nm, similar to those observed with metal porphyrins (Figure 26B, dotted line) (3, 22). Due to the strong absorbance of porphyrins, it is likely that all of the defining spectral characteristic of the Fe-S cluster have been obscured, except for the small feature visible at 340 nm in the aerobic sample, which potentially belongs to the Fe-S cluster (Figure 26B). Due to the spectral differences between aerobically and anaerobically purified enzymes we chose to perform all subsequent protein purifications and experiments under strict anaerobic conditions.

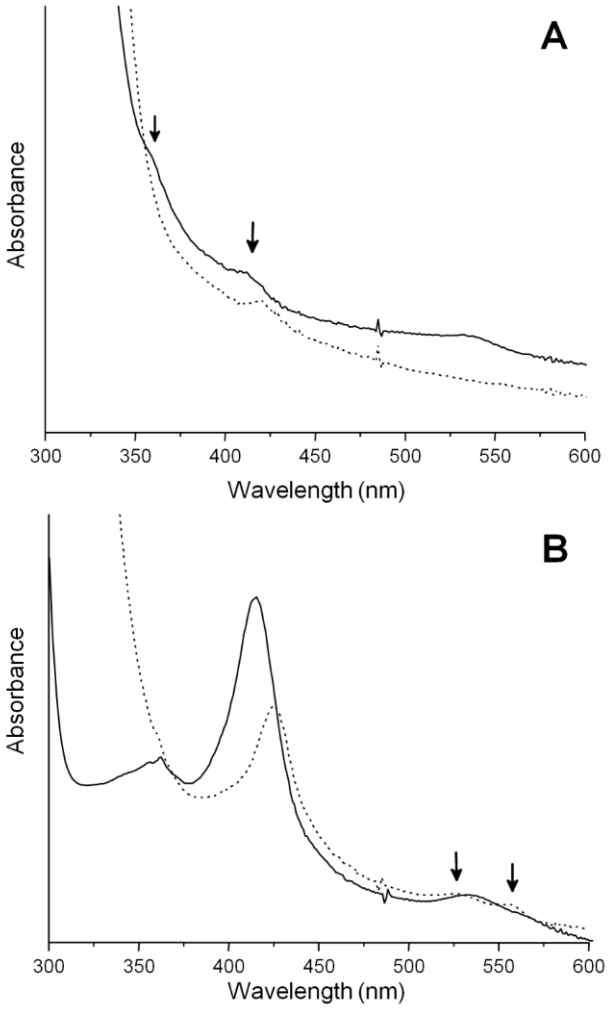

Figure 26.

UV/Vis characterization of BchE from *R. sphaeroides*. (A) UV/Vis trace of anaerobically purified BchE from *R. sphaeroides* (black trace) showing characteristic 4Fe-4S peaks at 350 nm and 420 nm (arrows) (31, 39). The same sample was made anaerobic and reduced by the addition of 1 mM dithionite (dashed line) shows a decrease in absorbance of the same features. **(B)** UV/Vis spectrum of aerobically purified BchE from *R. sphaeroides* (black) or reduced with dithionite (dashed line). Arrows indicate the emergence of twin peaks in the 500-600 nm region upon reduction suggestive of a metal porphyrin (22).

BchE interacts with MgPIX ME.

Upon successful purification of native BchE, we attempted to confirm its identity as the enzyme responsible for the conversion of MgPIX ME into DVP. To ascertain this, an anaerobic activity assay was carried out, containing BchE, partially purified MgPIX ME and PIX ME from *E.coli* cells expressing our heterologous pathway (22), with small

amounts of MgPIX also present in the reaction mixture. A mixture of porphyrins with and without magnesium was used, as it is possible that BchE is able to interact with both substrates, like several other BChl pathway enzymes (22). Regardless of incubation length, even up to 24 hours, no product formation was observed. However, a significant decrease of MgPIX ME, the predicted substrate for BchE, and some decrease in MgPIX and PIXME upon the addition of BchE to the reaction mixtures was evident (Figure 27, Figure 28). Neither SAM nor NADH was required in the assay mixture for the observed effect (Figure 27). Interestingly, replacement of NADH with equal volume of buffer in the reaction mixture was found to cause an increased loss of porphyrin substrates (Figure 27), suggesting that NADH is inhibitory. That result is only present in reactions containing BchE, as control conditions show the removal of NADH without the addition of enzyme has no affect on porphyrins present (Figure 29). While BchE contains a predicted cobalamin binding domain, additional or removal of adenosyl-cobalamin did not impact the results of our activity assays (Figure 30). We also attempted the experiment in the presence of hydroxy-cobalamin, as it is unknown which isoform of the cofactor BchE interacts with. However, the presence of hydroxyl-cobalamin in the activity assays made determination of metal porphyrins unfeasible due to an apparent aggregation of assay components (Figure 31). Finally, pre-incubation of BchE with a reducing agent, such as dithionite, or the subsequent addition of the reducing agent to the reaction mixture did not affect the activity of the enzyme (data not shown).

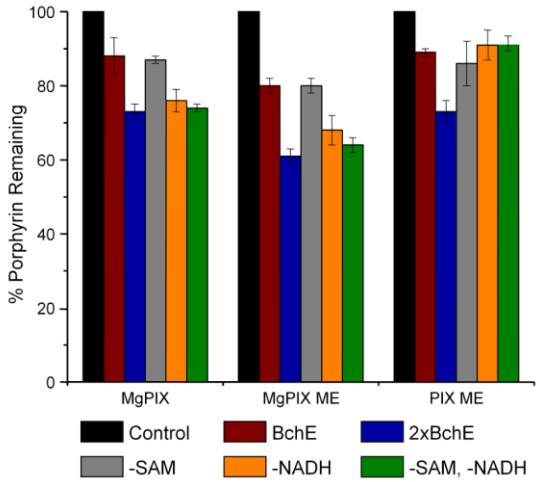

Figure 27.

Interaction of BchE with porphyrins under varying assay conditions. Numbers in the graph are reported as a percentage of control for each of the three porphyrins present in the reaction mixtures. Control samples all contained a mixture porphyrins, 500 µM SAM, 500 µM NADH and 10 µl of buffer. For samples BchE and 2xBchE, 5 or 10 µl (respectively) of freshly purified BchE from *R. sphaeroides* under anaerobic conditions was added. Upon addition of BchE, linear disappearance of porphyrins was observed from the reaction mixtures, as analyzed by HPLC. Interestingly, a higher percentage of MgPIX ME would disappear from the mixture compared to other porphyrins present. Subsequently, to determine if SAM, NADH or both were required for the observed effect, each was systematically removed from the reaction and replaced with equal volume of buffer. All reactions with removed cofactors contain 5 µl of BchE. Removal of NADH greatly increased the effect of BchE addition, while removal of SAM had no impact.

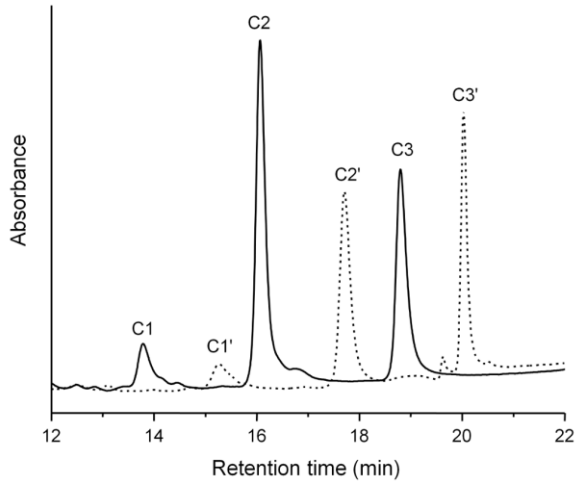

Figure 28.

HPLC analysis of BchE activity assay. HPLC analysis of activity assay mixtures containing either 10 µl of BchE purified from *R. sphaeroides* (dashed trace, correspond to 2xBchE in Table 1) or 10 µl of buffer in the control sample (solid trace). Peaks C1 and C1' correspond to magnesium protoporphyrin IX (MgPIX), peaks C2 and C2' correspond to magnesium protoprophyrin IX methylester (MgPIX ME) and peaks C3 and C3' correspond to protoporphyrin IX methylester (PIX ME). The amount of porphyrin present was measured by integration of the area under the corresponding peak,. It is worth noting that while the peak for C3' is taller that the peak for C3, the area under C3' is significantly lower, results are shown in Table 1. We have noticed the shift in the retention times during several individual experiments, and it is likely due to a failing HPLC column, however, it does not affect the separation of the porphyrin compounds. A wavelength of 412 nm was utilized for porphyrin detection.

Regardless of the composition of the reaction mixture, a higher percentage of MgPIX ME was consistently lost compared to both MgPIX and PIX ME, suggesting a preference of BchE towards MgPIX ME as a preferred substrate (Figure 27). As this is predicted to be a radical driven reaction, it is possible that during enzyme turnover a substrate-enzyme complex is being created, thus accounting for the disappearance of

substrate with no concomitant appearance of product. This theory is supported by the linear relationship between the amount of BchE added to the reaction mixture and the percentage of substrate missing during HPLC analysis compared to the control samples (Figure 27).

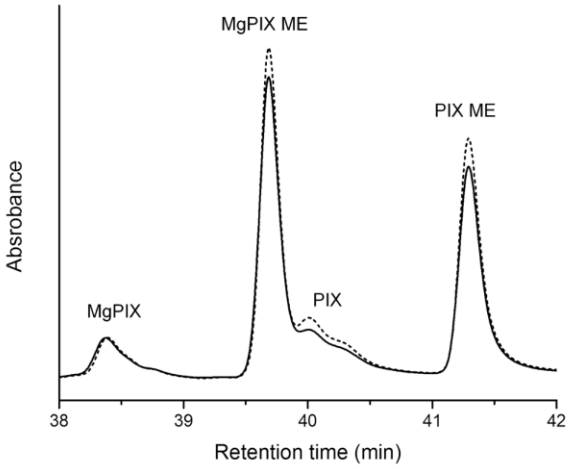

Figure 29.

Control BchE activity assays with and without NADH. Two identical control BchE activity assays were setup, one contained 500 µm NADH (solid trace) and in the second NADH was replaced with an equal volume of buffer (dashed trace). Removal of NADH from control reactions did not produce an effect like the removal of NADH from reactions containing BchE.

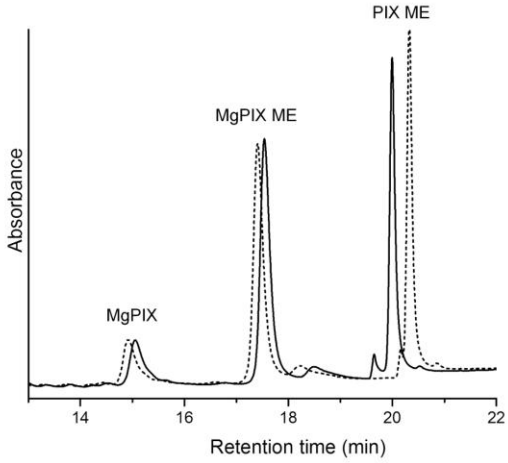

Figure 30.

Addition of adenosyl-cobalamine does not affect rate of porphyrin reduction in the BchE activity assays. Two BchE activity assays were setup under identical conditions with the exception that one condition contained 200 µM adenosyl-cobalamin (dashed trace) and the other was supplemented with an equal volume of buffer. The presence of adenosyl-cobalamin in the reaction mixture does not appear to have any effect on the rate of porphyrin disappearance upon addition of BchE.

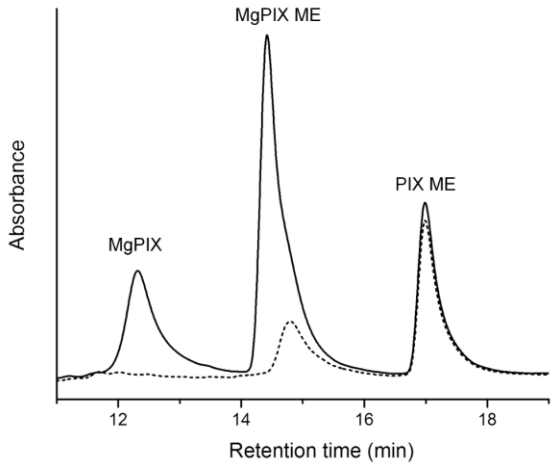

Figure 31.

Addition of hydroxy-cobalamine to control reactions depletes the metal porphyrins from the reaction. The addition of hydroxy-cobalamine (dashed trace) to the control mixture (solid trace) removes the majority of all metal porphyrins as detected by our HPLC assay. A large peak appeared at void volume, the composition of which could not be determined. It is possible that the missing porphyrins were part of the void peak.

Reconstitution and EPR characterization of BchE from *E. coli*.

Knowing that the Fe-S cluster is likely responsible for the initiation of the enzyme reaction, it was important to fully understand the cluster's composition, especially given that no activity of BchE purified from *R. sphaeroides* could be measured *in vitro*. One of the techniques well suited to characterize an Fe-S cluster is EPR, however, the caveat is that EPR requires a high protein concentration. Unfortunately, BchE purified from *R. sphaeroides* aggregates and precipitates prior to reaching the required concentration. Therefore, in order to circumvent these issues, we attempted to express and purify BchE from *E. coli*. Heterologous expression of BchE proved complicated as BchE containing a C-terminus 6xHis-tag would be degraded inside *E. coli* cells (Figure 24) and BchE containing a N-terminus 6xHis tag was only present in inclusion bodies, which required

solubilization in Urea. In vitro reconstitution of an Fe-S cluster following solubilization in Urea is possible and has proven effective for other 4Fe-4S enzymes (31).

Initial refolding and Fe-S reconstitution experiments were performed with full length $BchE_{1-612}$ purified from *E. coli* inclusion bodies and reconstituted under anaerobic conditions. While the UV/Vis measurements suggested the Fe-S cluster was successfully reconstituted, EPR signals of full length BchE were complex and noisy (Figure 32). Proteins containing a 4Fe-4S cluster have very characteristic EPR spectra, including an EPR silent 2^+ state upon initial purification, which gives way to a complex spectra upon the reduction of the 4Fe-4S cluster to the 1^+ state (31, 39, 42). When tested, reconstituted BchE gave an EPR signal after both the initial purification and after being reduced with dithionite (Figure 32). While the spectra from the reduced state was very similar to that of other 4Fe-4S proteins (31, 39, 42), we could not determine the source of the spectra from the 2^+ state following purification. BchE contains a cycteine rich region near the C-terminus (Figure 21) and it may be possible that these cysteines were forming an EPR active complex, resembling that of an Fe-S cluster, during reconstitution. To test this theory we cloned $BchE_{413-612}$ with an inclusion body localization tag (26), due to its otherwise poor expression and in order to maintain the same purification and reconstitution protocols as with other segments of BchE. Utilizing $BchE_{413-612}$ we were able to confirm that the last domain was contributing to the artifacts, complicating the EPR data for the reconstituted full length protein (Figure 33). Based upon sequence analysis and comparison between BchE homologues (Figure 19, Figure 21) we suggest

that the EPR signal produced by $BchE_{413-612}$ is the result of the reconstitution protocol and is not biologically relevant (Figure 32).

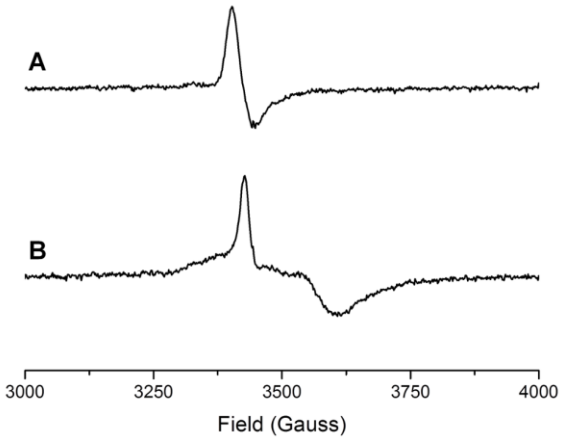

Figure 32.

EPR spectra of reconstituted $BchE_{1-612}$. Full length $BchE_{1-612}$ as purified from *E. coli* inclusion bodies and reconstituted under anaerobic conditions. **(A)** Spectrum of BchE immediately after reconstitution and purification under aerobic conditions and likely represents the oxidized state of any Fe-S clusters present. **(B)** Same protein sample reduced with 1 mM ditionite under anaerobic conditions. The spectrum in **A** is uncharacteristic of 4Fe-4S clusters which have been exposed to oxygen, as they should be EPR silent. However, the spectrum in **B**, after dithionite reduction is closer to what is expected from the 4Fe-4S cluster. The spectrum is an average of 6 scans obtained with the following parameters: temperature 20K, power 200 μW, receiver gain 60, modulation amplitude 10 G.

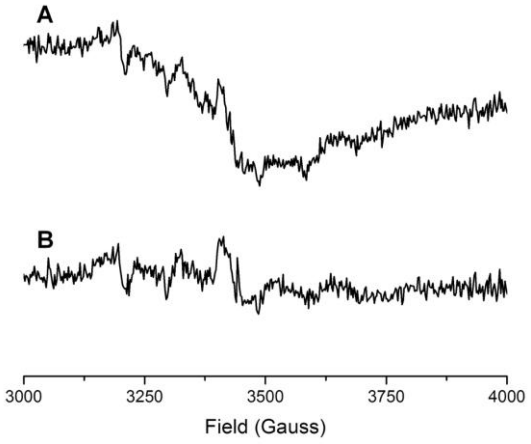

Figure 33.

EPR spectra of reconstituted $BchE_{413-612}$. (A) EPR spectrum of $BchE_{413-612}$ after reconstitution from *E. coli* inclusion bodies and aerobic purification. **(B)** Same sample after anaerobic reduction with 1 mM dithionite. The weak signal in **A** is due to the lower solubility of $BchE_{413-612}$ compared to other reconstituted segments of BchE, therefore it was difficult to concentrate the sample sufficiently to generate a stronger EPR signal. The spectrum is an average of 6 scans obtained with the following parameters: temperature 20K, power 200 μW, receiver gain 60, modulation amplitude 10 G.

To avoid the background noise from the non-specific Fe-S insertion arising during reconstitution of full length BchE, we shifted our attention to $BchE_{192-393}$, which is predicted to contain the highly conserved 4Fe-4S cluster found in radical SAM enzymes (Figure 20). $BchE_{192-393}$ was similarly overexpressed and found to be present only in inclusion bodies in *E. coli*. Following solubilization and reconstitution, $BchE_{192-393}$ was redox active and displayed a dark brown color, typical of Fe-S enzymes (31, 39). The

spectral features and redox activity of refolded $BchE_{192-393}$ were similar to those of BchE purified from *R. sphaeroides* (Figure 34). Reconstituted $BchE_{192-393}$ had absorption features with maxima at 330 nm and 415 nm, suggesting the presence of a 4Fe-4S cluster (Figure 34) (31, 39, 42, 43). The characteristic decrease in absorbance was present in both native and reconstituted proteins upon reduction with dithionite under anaerobic conditions (Figure 34A). The slight shift in the spectral features was attributed to the Fe-S clusters between the reconstituted $BchE_{192-393}$ and BchE purified from *R. sphaeroides*. Slight changes in the local environment near the Fe-S cluster, such as water accessibility, may account for this slight shift of absorbance features between the two samples (Figure 26A, Figure 7).

Given the similarity observed in the redox features of the reconstituted $BchE_{192-393}$ compared to the native BchE, and that the reconstituted $BchE_{192-393}$ was soluble at concentrations needed for EPR, $BchE_{192-393}$ was utilized for all subsequent experiments. Upon reduction of $BchE_{192-393}$ to the 1+ state a distinct EPR feature characteristic of 4Fe-4S clusters was apparent (Figure 35A). The complex feature at g=1.94 of the reduced state is in good agreement with previously published data for other 4Fe-4S proteins, both reconstituted and purified from a native host (31, 39, 42, 43). This data further confirms unequivocally that BchE contains a redox active 4Fe-4S cluster which may be responsible for initializing the cyclization of MgPIX ME to DVP.

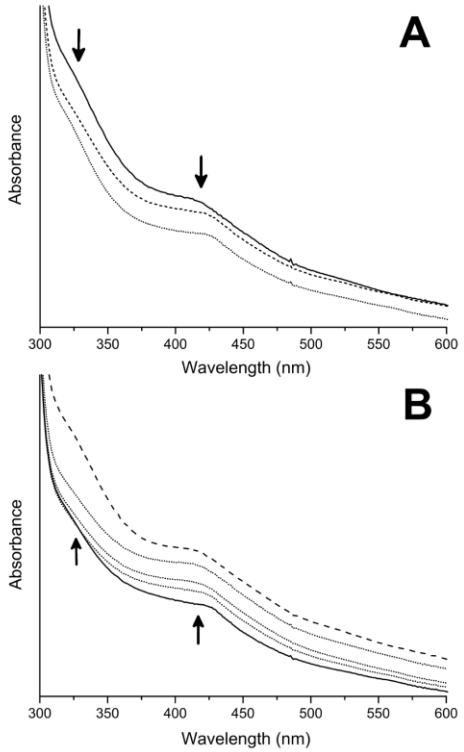

Figure 34.

Redox activity of reconstituted BchE. (A) UV/Vis analysis of reconstituted $BchE_{192\text{-}393}$ as isolated (black line), or with increasing concentrations of dithionite in 100 µM steps up to 1 mM final concentration (dashed and dotted lines). Arrows indicate the location and the direction of change of the major 4Fe-4S associated features at 340 nm and 420 nm (31, 39). **(B)** Re-oxidation of reconstituted $BchE_{192\text{-}393}$ by the addition of SAM. The trace shows the "as purified" sample (dashed line), the sample after reduction with dithionite (solid line) and a series 100 µM additions of SAM up to 1 mM final concentration (dotted lines). Arrows indicate the direction of the absorbance change upon addition of SAM.

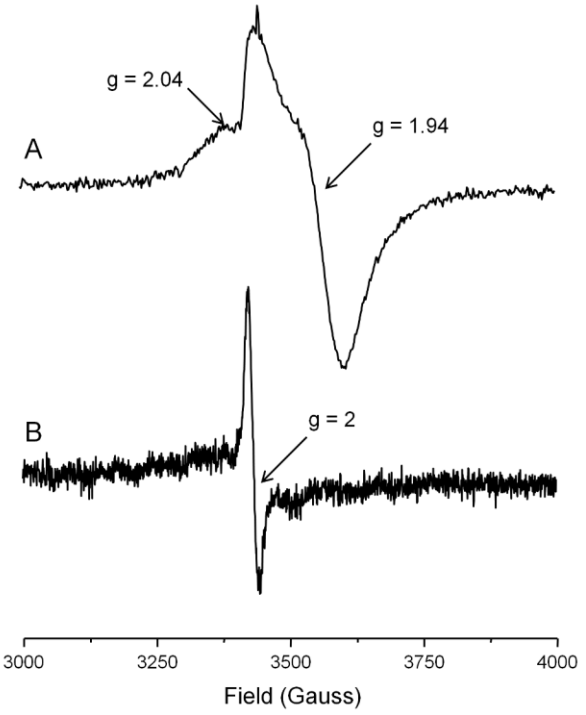

Figure 35.

EPR of BchE. Reconstituted $BchE_{192\text{-}393}$ was purified in sufficient concentration to create an EPR sample. **(A)** EPR signal of $BchE_{192\text{-}393}$ reduced with 1mM dithionite under strictly anaerobic conditions. **(B)** Same sample used to generate trace A, mixed with 2 mM SAM and allowed to incubate prior to being frozen. Note the drastic shift from complex EPR spectrum of the 4Fe-4S with g=1.94 to a spectrum with a g=2 between upon the addition of SAM. The spectrum is an average of 6 scans obtained with the following parameters: temperature 20K, power 200 µW, receiver gain 60, modulation amplitude 10 G.

Interaction of BchE with SAM.

The initial step in the conversion of MgPIX ME to DVP is hypothesized to require a transfer of an electron from a reduced (1^{+}) Fe-S cluster to a SAM molecule (33). If this electron transfer occurs, it can be detected by monitoring the changes in the

UV/Vis as well as the EPR spectrums attributed to the Fe-S cluster. We first monitored this change produced by the addition of SAM to the reduced, anaerobic $BchE_{192-393}$ over the UV/Vis region (Figure 34). As expected, upon interaction of SAM with $BchE_{192-393}$, the UV/Vis spectrum returned to pre-reduction levels (Figure 34).

Similarly, the EPR signal from the reduced 4Fe-4S cluster should decrease, as the cluster transitions from the EPR active 1^+ to the EPR silent 2^+ state upon the addition of SAM. Furthermore, the addition of SAM to a reduced radical SAM enzyme will not only to quench the 4Fe-4S EPR signal, but to also may give rise to a sharp EPR signal at g=2, characteristic of an organic radical. This g=2 signal likely results from a 5'-adenosyl radical, formed during the interaction of SAM with the reduced Fe-S cluster (39). To investigate this, we took the reduced, anaerobic $BchE_{192-393}$ (Figure 35A), added SAM under strictly anaerobic conditions and re-froze the sample following a 30 minute incubation. Subsequently, the sample was analyzed by EPR (Figure 35B). The primary feature with g=1.94 associated with the 1^+ state of the Fe-S cluster present in the reduced sample disappears and is replaced by a sharp feature at g=2 (Figure 35B). The location and shape of the feature present after the addition of SAM is similar to that previously suggested to result from the formation of a 5'-adenosyl radical (39), therefore it could be that BchE uses SAM to form the initializing radical for the cyclization of MgPIX ME to DVP.

DISCUSSION

The cyclization reaction of MgPIX ME to form DVP remains the least well understood enzymatic step involved in the formation of BChl. (12, 20, 44, 45) To date,

the predicted reaction mechanism has not been confirmed biochemically. Studies have indicated that the source of the oxygen on the keto group of the created E ring differs between AcsF and BchE (19), however, the mechanistic details are still not clear, owing to the fact that the cyclase is a difficult enzyme to express. AcsF's likely requirement for at least two poorly understood components for function complicates studies into its reaction, as does the fact that AcsF itself is membrane associated. BchE is similarly complicated, requiring at least two cofactors and potentially at least one other, as of yet unidentified component for function.

Our initial aim was to study the activity of both putative MgPIX ME cyclases in the context of a robust and well characterized system for overproduction of magnesium porphyrins in *E. coli*. Unfortunately, neither enzyme proved to have any detectable *in vivo* activity, likely because both proteins proved difficult to functionally express. In fact, we were unable to detect any significant levels of AcsF in our expression trials, which is perhaps not surprising given that some membrane associated diiron oxygenases are known to be poorly expressed in *E. coli* (46). In some instances, overexpression of diiron oxygenases required substantial optimization of growth and vector conditions for successful expression (46). For this reason, we chose not to pursue further work with AcsF and instead focused on BchE, which we were able to stably express in *E. coli*.

Some clues to the potential maturation requirements of BchE were gleamed from the purification of the GroEL chaperone machinery during Ni^{+} affinity chromatography of BchE from *R. sphaeroides*. The interaction of BchE and the GroEL chaperone complex is not entirely surprising as BchE is predicted to contain several α-β domains,

which have been identified as substrates for the GroEL folding machinery (47). GroEL is also known to interact with proteins which contain large regions of hydrophobicity and have a propensity to aggregate (47-49). Based on this information, we postulated that BchE may have a requirement for GroEL in order to fold correctly. Interestingly, coexpression of GroEL chaperones from a plasmid-based system and BchE in *E. coli* did not produce any improvement in the yield of soluble protein in our hands.

Understanding how to optimize expression of the functional enzyme was important, and required a detailed understanding of the catalytic activity of BchE. The novel findings reported here bring us closer to that goal. The redox and EPR data from purified native BchE as well as reconstituted $BchE_{192\text{-}393}$ confirm that BchE is a 4Fe-4S radical SAM enzyme. The generated EPR spectrum is highly similar to other known 4Fe-4S radical SAM enzymes (31, 39, 42, 43). Even more intriguing is the quenching of the Fe-S EPR signal and the rise of a radical like g=2 feature upon the addition of SAM to the reduced enzyme. A similar g=2 feature was observed upon mixing of another putative 4Fe-4S enzyme with SAM and was suggested to belong to the 5'-deoxyadenosyl radical generated upon SAM cleavage (39). The oxidation of the BchE 4Fe-4S cluster by SAM confirms the ability of BchE to transfer electrons onto SAM molecules. This electron transfer is hypothesized to be the initial step during the conversion of MgPIX ME to DVP (33). It is still unclear how many SAM molecules are utilized during a single turnover of BchE, though, it has been suggested that two SAM molecules are cleaved per reaction (33).

Given that the data strongly suggests that the 4Fe-4S cluster is redox active and is able to interact with SAM, it is still unclear why BchE purified from *R. sphaeroides* is not able to perform a complete enzymatic turnover or why the removal of SAM from activity assays has no effect on the disappearance of porphyrins. Similarly, it is not clear why we did not find cobalamin to be essential for the activity of BchE, even though it is known to be important for *in vivo* activity (21). As multiple isotypes of cobalamin exist, it is possible that we have simply not tried the correct one during our work. Conversely, removal of NADH from reaction mixtures greatly increases porphyrin loss upon the addition of BchE. It is possible that the interaction of SAM and BchE is strong enough that BchE from *R. sphaeroides* co-purifies with SAM and thus does not require addition of SAM to the reaction mixture. Furthermore, since NADH and SAM both contain an adenosyl moiety, there is potential for NADH to replace one of the SAM molecules near the active site and inhibit BchE activity, explaining higher porphyrin loss in the absence of NADH from reaction mixtures. Interestingly, recently characterized HemN, also a radical SAM enzyme, is only functional in the presence of a crude cell lysate and requires the presence of NADH (50). The missing component required for HemN activity is yet to be identified. Similarly, another protein or co-factor is likely essential for successful catalytic turnover of BchE. Although in our hands, addition of *R. sphaeroides* crude cell lysate to the reaction mixture did not promote BchE activity.

While HemN is the closest known *E. coli* homologue of BchE, it is suggested that BchE actually belong to the p-methylase protein subfamily (33, 51, 52). It is likely that the reaction is initiated through the cleavage of SAM by the reduced 4Fe-4S cluster,

generating a 5'-adenosyl radical. It is suggested that the radical is then transferred from 5'-adenosyl to the substrate, which in turn may interact with the cobalamin and abstracts a hydroxyl group (33). While this theory explains the introduction of the oxygen onto what will become the E ring of DVP, it does not suggest a mechanism by which the actual cyclization of $C13^2$ to C15 occurs. An interesting possibility because of the complexity of this reaction and the high reactivity of radical species, is that a chaperone could be involved to guide the radical chemistry. For example, protein guides have been suggested in the case of lignan biosynthesis in order to help select for desired stereochemistry (53). A similar chaperone could be envisioned in the case of BchE.

The co-purification of BchE from *R. sphaeroides* with what appears to be a porphyrin molecule under aerobic conditions suggests that oxygen may interfere or disrupt the BchE catalytic cycle. Since BchE utilizes a radical based mechanism, it is not unreasonable that the presence of oxygen could disrupt the normal reaction cycle and perhaps generate dead end protein-substrate complexes. This mechanism could ultimately explain the observed decrease of substrate but a lack of product formation during our activity assays, as we were unable to quench the reactions inside of the anaerobic chamber. This also supports a case for a potential chaperone protein that could guide the radical reaction, ensuring complete cyclization and preventing covalent modification of the enzyme by the reaction intermediate.

To our knowledge, this is the first published study attempting to demonstrate the function of both AcsF and BchE as parts of BChl biosynthetic pathway in a heterologous host. Furthermore, our work confirms for the first time that BchE from *R. sphaeroides* is

indeed a member of the radical SAM enzyme family and contains a 4Fe-4S cluster, as predicted from sequence analysis. Additionally, this is first time BchE was purified from the native host and shown to interact with intermediates of the BChl biosynthetic pathway. While there is still significant work to be done before the process of cyclization of MgPIX ME to DVP is fully understood, this work serves as the establishment of methods required to understand this complex reaction. Overall, these developments bring us significantly closer to understanding the last uncharacterized enzyme from BChl biosynthesis, an enzyme responsible for the formation of a highly unique fifth ring on the tetrapyrrole backbone. The understanding of this key step in BChl biosynthesis also brings us significantly closer to functionally extending the pathway in a heterologous host.

Chapter 5. Heterologous expression of the bacterial reaction center from *Rhodobacter sphaeroides* in *E. coli.*

Chapter 5 Overview

Photosynthetic organisms are able to capture sunlight and convert it into chemical energy, which is used by the organism for growth. Nature has evolved several mechanisms to enable this conversion with plants algae utilizing expansive enzyme complexes. Many other bacteria, such as *Rhodobacter sphaeroides* use a simpler system with only three core subunits, termed the bacterial reaction center. The relative simplicity of the bacterial reaction center makes it an attractive option for anyone wanting to bring the light-harvesting capabilities to a nonphotosynthetic host. In this work we explore heterologous expression and purification of the reaction center from *R. sphaeroides* in *E. coli*, as a model organism. Furthermore, we demonstrate utilizing fluorescent fusions that the reaction center complex is only stably present upon the expression of all three subunits and is primarily found at the poles of the *E. coli* hosts. These discoveries help bring us significantly closer to the functional expression of the reaction center and hence the ability for light-energy conversion in a nonphotosynthetic host.

INTRODUCTION

Throughout the work presented in this thesis, the underlying theme has been the reconstruction of biochemical pathways in order to engineer light-energy conversion into a non-photosynthetic host. In addition to the production of biosynthetic pigments, as was discussed earlier in this thesis, a photosynthetic reaction center (RC) is needed in order to

capture and convert light into energy. For this task, we have chosen to work with the bacterial reaction center from *R. sphaeroides* as it is the model RCs and has been extensively characterized (1-3).

R. sphaeroides utilizes a type II reaction center, which means that it is anoxygenic and the excited electrons are cycled back to the reaction center using a cytochrome c-2 (4-6). The RC itself is composed of two transmembrane subunits, *pufM* and *pufL*, as well as a primarily cytoplasmic subunit, *puhA* (referred to as M, L and H subunits, respectively) (Figure 36). Functional RCs present in *R. sphaeroides* are known to bind several cofactors to form an electron transfer chain (5). Specifically, two bacteriochlorophylls found near the periplasmic side of the complex compose what is referred to as the "special pair", the source of high energy electrons upon light excitation (6, 7). After excitation, the electrons are subsequently transferred through another bacteriochlorophyll, a bacteriopheophytin and a pair of quinones with ubiquinone serving as the final electron acceptor (6, 7). In addition to the cofactors required for electron transfer, a carotenoid and several lipid molecules are known to interact with the RC (Figure 36)(8-11). The location and confirmation of the cofactors within the RC complex is now well established due to a multitude of available crystal structures (4, 12-15).

After two excitation reactions, a pair of high energy electrons is transferred to the final quinone acceptor. The fully reduced quinone is then released into the membrane space (7, 16, 17). From there, the electrons are shuttled through a redox chain, with each step in the chain coupled to the translocation of protons into the periplasm, thus creating a proton gradient as the means of converting light into chemical energy (18, 19). This

proton gradient can be used by the cell in order to regenerate ATP or reduce NAD(P) to NAD(P)H. In an engineered host, the supplemented pools of high energy molecules can be funneled into complex biosynthetic pathways, such as production of biofuels (Figure 3) (20, 21). The ability to supplement the nonphotosynthetic host's energy pools with the functional expression of the RC could relieve some of the metabolic burden created from complex biosynthetic pathways.

While protocols for working with and purifying RC from *Rhodobacter* are now well established, little has been done regarding heterologous RC expression in a nonphotosynthetic host (22). A single attempt has been published regarding expression and purification of individual RC subunits from *E. coli* (3) and nothing exists in the literature about functional expression of the RC complex. In the case of expression of individual subunits, the expression proved to be poor and over 10 L of cells were required for purification of each subunit (3). This perhaps should not be surprising as the RC complex is composed of two membrane protein, and membrane proteins are notoriously difficult to overexpress in a heterologous host, frequently causing toxicity issues (23). We believe that by gaining a better understanding of how the RC behaves when expressed in *E. coli*, it is possible to minimize the toxicity problems previously reported. Furthermore, with optimization, it should be possible to not only express the RC but to also tie its proton translocation ability into the host's metabolism.

In this work we report the first successful expression of the complete RC complex in *E. coli* as well as the first partial purification of the complete RC assembly from a heterologous host. Furthermore, we explore the localization and stability of the RC within

E. coli. Together, these observations bring us significantly closer to the final goal of the project, using light-energy conversion as a means to supplement a nonphotosynthetic host's energy pool.

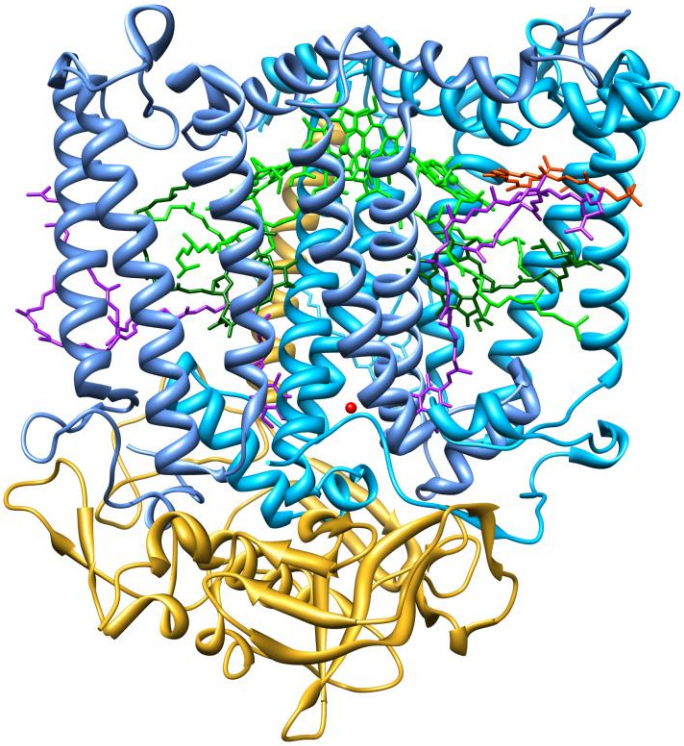

Figure 36.

Crystal structure of the RC from *R. sphaeroides* solved to 2A (PDB 3I4D) and presented with the cytoplasmic H subunit oriented downward. Transmembrane subunits are shown in blue with L (dark blue) and M (light blue). The H subunit is shown in gold. Bacteriochlorophyll molecules are shown in light green and bacteriopheophytin is shown in dark green. Iron is represented by a red sphere. Bound spheroidene molecule is shown in orange and the ubiquinone is represented in purple. (Fujii *et al.* Unpublished)

MATERIALS AND METHODS

Enzymes and Chemicals:

Chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO), unless otherwise described. Restriction enzymes and DNA polymerases were purchased from New England Biolabs (Ipswish, MA).

Cell growth conditions:

E. coli JM109(DE3) was utilized for DNA manipulation and protein expression. Bacterial cultures were grown at 37°C or as described in the text, in Luria-Bertani (LB) medium supplemented with 30 µg/ml of kanamycin, 100 µg/ml ampicillin or 50 µg/ml chloramphenicol as needed for plasmid maintenance.

Plasmid construction:

In order to generate plasmids pBBRBB-Mhis, pBBRBB-L and pBBRBB-H the *pufM* (PMID 3719398), *pufL* (PMID 3719399) and *puhA* (PMID 3719203) genes were amplified using primer pairs P1 and P2, P3 and P4, P5 and P6, respectively (Table 7) from the *R. sphaeroides* 2.4.1 genomic DNA isolated with Wizard Genomic DNA purification kit (Promega). The pBBRBB-GFP vector (24) and the generated PCR products were then digested with *Bgl*II and *Not*I and ligated together in order to generate reaction center expression vectors. The C-term 6x-Histidine tags were incorporated into the amplification primers (Table 7). Assembly of multi-gene constructs was performed using standardized BioBrick™ cloning cut sites (24). Briefly, the vectors were opened with either *Eco*RI and *Xba*I or *Pst*I and *Spe*I, while the gene to be inserted was cut with *Eco*RI and *Spe*I or *Pst*I and *Xba*I, the vector and insert combination containing compatible sticky ends were sequentially ligated together, see Table 8 for final gene arrangement.

Fluorescent proteins fusions were made by cloning the amplified gene of interest in frame into the *Bgl*II and *Nde*I sites of either pACBB-ntH6-GFP or pACBB-ntH6-DsRed (24). Reaction center genes were amplified with the primer pair P7 and P8 for *pufM*, P9 and P10 for *pufL* and P11 and P12 for puhA (Table 1). The PCR products were digested with *Bgl*II and *Nde*I and cloned into the same sites of either pACBB-ntH6-DsRed or pACBB-ntH6-GFP to generate pACBB-*pufM-Dsred*, pACBB-*pufL-GFP* and pACBB-*puhA-GFP*. Gene stacking was performed as described above. Plasmids used in this study are listed in Table 8.

Table 7. Primer utilized in Chapter 5.

P1 (pufM F)	5'-agtAGATCTcgATGGCTGAGTATCAGAACATCTTC-3'
P2 (puMHis R)	5'-atgcggccgcTCAatgatggtgatgatggtgGTTCAGCGGCGCCAT-3'
P3 (pufL F)	5'-GcTccAGATCTTCCGatgGCACTGCTCAGCTTC-3'
P4 (pufL R)	5'-atcgagcggccgcttaTCAGCCATTGATGCCTC-3'
P5 (puhA F)	5'-GcTccAGATCTTCCGatgGTTGGTGTGACTGCTTTTG-3'
P6 (puhA R)	5'-atcgagcggccgcttaGGCGTATTCGGCCAGC-3'
P7	5'-acagatctATGGCTGAGTATCAGAACATCTTCTCCC-3'
P8	5'-gaCATATGagaaccagaaccagaaccagAGTTCAGCGGCGCCATGCC-3'
P9	5'-acagatctATGGCACTGCTCAGCTTCGAGC-3'
P10	5'-gaCATATGagaaccagaaccagaaccagAGCCATTGATGCCTCCCG-3'
P11	5'-acagatctATGGTTGGTGTGACTGCTTTTGGAAAC-3'
P12	5'-gaCATATGagaaccagaaccagaaccagAGGCGTATTCGGCCAGCATCGC-2'

Protein Expression and Purification:

E. coli JM109(DE3) was utilized for heterologous expression of reaction center components. For purification, a 4 ml overnight culture started directly from a plasmid transformation containing with the desired combination of the reaction center subunits in *E. coli* JM109(DE3), was transferred into 1 L of LB and allowed to grow for 12 hours at 25°C prior to harvest by centrifugation. Harvested cells were resuspended in 20 ml of buffer A (50 mM Tris-HCl pH 8, 250 mM NaCl, 5 mM imidazole) and lysed by sonication. After lysis, N-N-dimethyldodecylamine-N-oxide (LDAO) was added to the

final concentration of 0.75% and allowed to incubate at 4°C for 2 hours, with gentle mixing. Subsequently, unsolubilized membranes and cell debris were removed by centrifugation at 12,000 rpm for 30 min at 4°C. The soluble lysate was passed over a 1 ml HisTrap FF Crude column (GE Lifesciences) equilibrated with buffer A plus 0.1% LDAO. Bound reaction center was eluted from the column via gradient of 0-100% buffer B (50 mM Tris-HCl pH8, 250 mM NaCl, 250 mM imidazole, 0.1% LDAO). Elution location of the reaction center components was analyzed via SDS-PAGE followed by western blotting using anti-his antibodies (R&D Systems) for complexes containing the 6xHis tag. In the case of GFP fusions, elution location was determined by monitoring fraction fluorescence at 525 nm with the excitation wavelength of 475 nm.

Fluorescent and Light Microscopy:

Cultures containing fusions of reaction center subunits and fluorescent proteins were grown at 25°C in LB and appropriate antibiotic until the OD600 of approximately 0.6. At the desired OD600, 100 ul of culture was centrifuged, washed twice with PBS and resuspended in 100 ul of PBS prior to use. Subsequently, 10 ul of PBS washed culture was placed on a poly-L-lysine coated microscope slide and allowed to attach for 30 min prior to light/fluorescent microscopy.

A Nikon Eclipse E800 photomicroscope equipped with bright field, Differencial Interference Contrast (DIC) and fluorescent optics was used coupled to Roper CoolSnap HQ monochrome camera capable of 16-bit digital images. The samples were viewed using a 100x, 1.4 n.a. plan apo objective. For fluorescent images, blue (excitation filter

470-490 nm, barrier 520-580 nm) and green (excitation filter 510-560nm, barrier 570-620nm) filters were used to detect the presence of eGFP and DsRed, respectively.

Table 8. Plasmids created in Chapter 5.

Plasmid name	Notes
pBBRBB-*pufM*$_{His}$	Contains *R. sphaeroides pufM* with a C-terminus 6xHis tag
pBBRBB-*pufL*	Contains *R. sphaeroides pufL*
pBBRBB-*puhA*	Contains *R. sphaeroides puhA*
pBBRBB-*pufL*$_{His}$	Contains *R. sphaeroides pufL* with a C-terminus 6xHis tag
pBBRBB-*puhA*$_{His}$	Contains *R. sphaeroides puhA* with a C-terminus 6xHis tag
pBBRBB-*pufM*$_{His}$-*pufL*	Contains *R. sphaeroides pufM* with a C-terminus 6xHis tag and *pufL*
pBBRBB-*pufM*$_{His}$-*pufL*-*puhA*	Contains *R. sphaeroides pufM* with a C-terminus 6xHis tag and *pufL* and *puhA*
pBBRBB-*pufM*$_{His}$-*pufL*$_{His}$	Contains *R. sphaeroides pufM* with a C-terminus 6xHis tag and *pufL* with a C-terminus 6xHis tag
pBBRBB-*pufM*$_{His}$-*pufL*$_{His}$-*puhA*$_{His}$	Also referred to as pBBRBB-RChis
pACBB-M-DsRed	Contains *R. sphaeroides pufM* with a C-terminus fusion to DsRed
pACBB-L-GFP	Contains *R. sphaeroides pufL* with a C-terminus fusion to GFP
pACBB-H-GFP	Contains *R. sphaeroides puhA* with a C-terminus fusion to GFP
pACBB-*pufM*$_{His}$-*pufL*$_{His}$-H-GFP	Contains *R. sphaeroides pufM* and *pufL* with C-terminus 6xHis tag in addition to *puhA*-GFP fusion

RESULTS

Expression and purification of the RC complex in *E. coli*.

When purifying the RC complex from *R. sphaeroides*, the preferred method for isolation of the entire complex is utilizing metal affinity chromatography coupled to a single 6xHis tag on PufM (22). Similarly, in our initial expression trials only the M subunit contained a 6xHis C-terminus tag, which corresponded to a single band visible during the development of an anti-his western blot employing total cell lysate (Figure 37). During the expression trials, high instability of the reaction center subunits was observed. Lowering growth temperatures to 25°C provided the best compromise between cell growth and the stability of the reaction center complex (Figure 38). RC stability was also tested among multiple *E. coli* strains, with *E. coli* JM109 (DE3) yielding the best

results (data not shown). Furthermore, significant variation of RC levels was noted between individual colonies picked from the same plate. Ultimately, to confirm the expression of all three subunits, 6xHis C-terminus tags were applied to M, L and H subunits creating the RChis construct. As expected, Western blotting following SDS-PAGE revealed the presence of three distinct bands, when analyzing cell lysates from RChis cells (Figure 37).

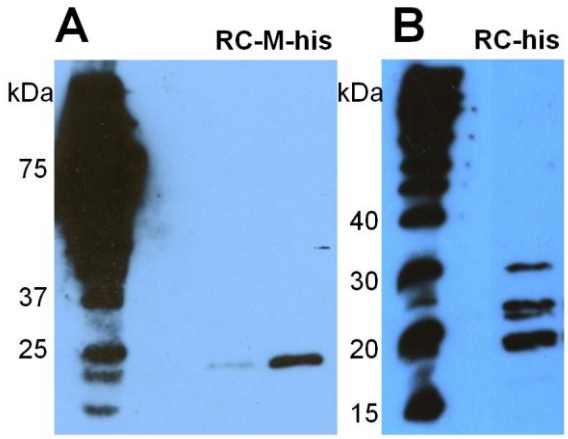

Figure 37.

RC center expression in *E. coli*. The full RC complex was expressed in *E. coli* with either a single 6xHis-tag on the M subunit **(A)** or 6xHis-tags on L, M and H subunits **(B)** and analyzed by Western blot with anti-his antibodies. Molecular weight ladders are present on the left side of each gel and molecular weights corresponding to each band are shown. Expected sizes for the L, M and H subunits are 21 kDa, 24 kDa and 28 kDa, respectively.

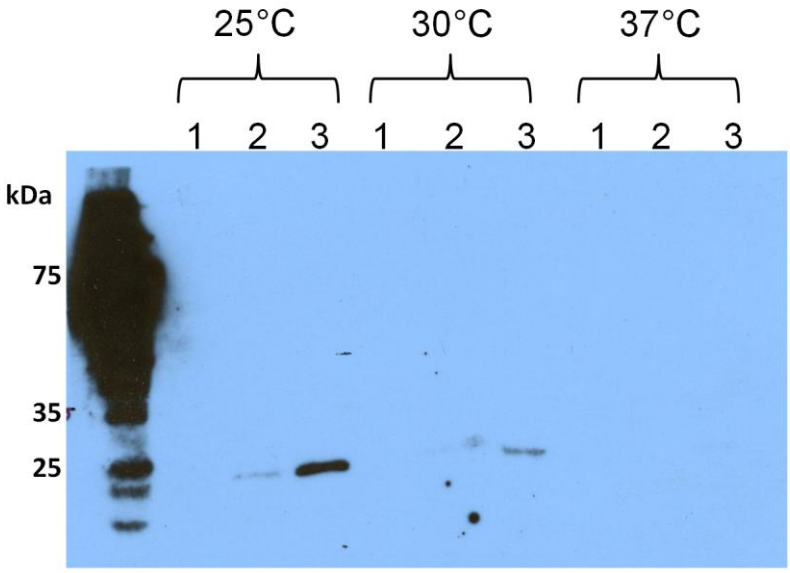

Figure 38.

Dependence of RC expression level on culture temperature. Three colonies from the same transformation of pBBRBB-RChis were each resuspended in distilled water. Each cell suspension was used to inoculate three 4 ml cultures of LB, which were subsequently grown at either 25°C, 30°C or 37°C overnight. Equal volume samples were taken from each of the cultures and the cell lysate was analyzed by anti-His Western blot. Lanes labeled 1, 2 or 3 display the original colony used. Temperatures marked above the gel represent growth temperature for that set of cultures.

After achieving stable RC expression, we attempted to purify the protein complex using a single 6xHis-tag located on PufM, as is commonly performed in *R. sphaeroides* (22). The presence of the 6xHis-tag allowed for the detection of PufM in the eluting fractions via western blot. The RC is known to elute from the Ni^{+} column relatively easily, at approximately 50 mM imidazole (22), which does not yield pure protein. In an attempt to improve the binding of the RC to the purification resin, we repeated the purification protocol with the RChis construct. The purification of RChis resulted in the elution of all of the subunits from the Ni^{+} column simultaneously; unfortunately, the proteins still eluted around 50 mM imidazole (Figure 39).

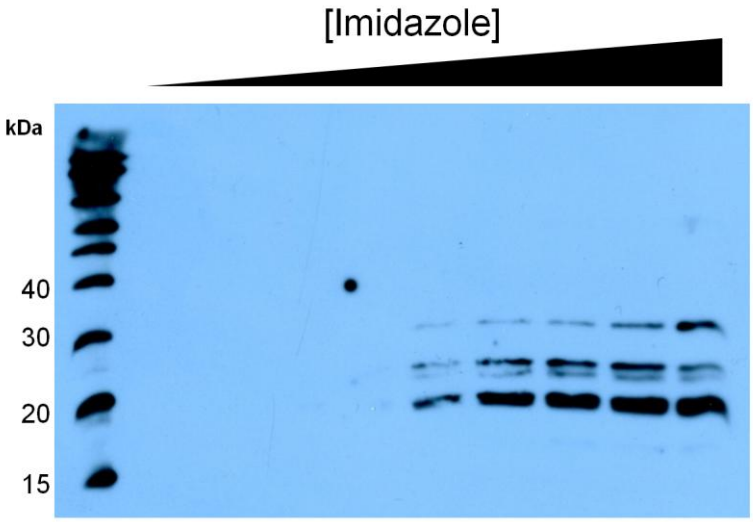

Figure 39.

Purification of RC-His from *E. coli*. Purification of the RC-His complex under increasing imidazole concentrations from the Ni^{+} resin as shown by the anti-His western blot. Lanes going from left to right on the gel show elution from the column as increasing imidazole concentrations are applied to the resin. Molecular weights of the protein ladder are shown. Imidazole concentration was increased from 20 mM to 50 mM from left to right.

Given the early elution and the amount of contaminants eluting along with the RC, the only way to confirm the presence of RC was by western blot. Given that western blotting is a time intensive technique, we wanted a simpler reporter of the presence of the RC complex. In order to monitor the elution in real time, a fusion between PuhA, the primarily cytoplasmic subunit, and eGFP was made. PuhA was chosen for the GFP fusion as we assumed a complex between PufM and PufL was required for the interaction of either subunit with PuhA (25). Under this design, so it should be possible to track the fluorescent of from the PuhA-GFP fusion as a sign of reaction center complex stability. Gradient elution of a Ni^{+} column loaded with the RC complex solubilized from *E. coli* cells revealed that a GFP signal associated with PuhA-GFP was eluting off the column at approximately the same time as PufM and PufL, as identified by western blot. This further demonstrated that the RC proteins eluted as a stable complex during purifications

and that tracking GFP was a viable alternative to Western blotting to follow RC elution. Furthermore, these results confirm not only successful expression of the RC complex in *E. coli*, but also our ability to isolate the RC as an intact complex.

Localization of Reaction Center proteins in *E. coli.*

During RC purification attempts, we repeatedly noted the toxicity to the host associated with RC expression. While toxicity is frequently associated with membrane protein expression, we wanted to see if we could glean any insights specific to the RC expression. Upon literature review, we noted the RC co-purified with cardiolipin, which is present at lower levels and has different localization in *E. coli* compared to the *R. sphaeroides* (9, 10). Cardiolipin is a phospholipid that has been associated with curvatures in the cell membrane and in *E. coli*, and is primarily found at the poles of the cell (26, 27). If the RC requires cardiolipin for stability, it is possible that the complex can only successfully form at the poles of *E. coli* hosts, thus further aggravating the toxicity already associated with membrane protein expression.

To determine the localization of individual RC subunits as well as the complete RC complex, fusions of individual subunits to fluorescent proteins were created. The fluorescent fusions were introduced into *E. coli* cells either on their own or in tandem with a second vector encoding the complete RC complex (Figure 40). We initially looked at the localization of the H subunits, as it is primarily cytoplasmic and only contains a single transmembrane helix (14, 15). Upon expression of just the H-GFP fusion in the cells, we noted aggregation in several parts of the cell in addition to some diffuse fluorescence (Figure 5A). Upon coexpression of H-GFP with the complete RC complex,

the fluorescence became localized at the cell's poles (Figure 5B). A similar pattern was also observed when looking at the localization of M-DsRed fusion. When expressed alone, M-DsRed created a patchy fluorescence pattern similar to that of H-GFP (Figure 5C). The fluorescence shifted to primarily polar localization upon co-expression of H-GFD with the rest of the RC complex (Figure 5D). Interestingly, little difference was observed upon expression of L-GFP alone (Figure 5E) or with the RC complex (Figure 5F). In both cases, the fluorescence associated with the expression of L-GFP fusion was localized throughout the cell, with little of the protein shifting to the cell's poles upon addition of the RC complex. Expression of just the fluorescent proteins with and without the presence of the RC complex resulted in even, diffuse fluorescence in the cell (GFP control shown in Figure 5G). The known polar localization of cardiolipin in *E. coli* coupled to our localization experiments suggests that cardiolipin might play an important role for RC assembly or stability.

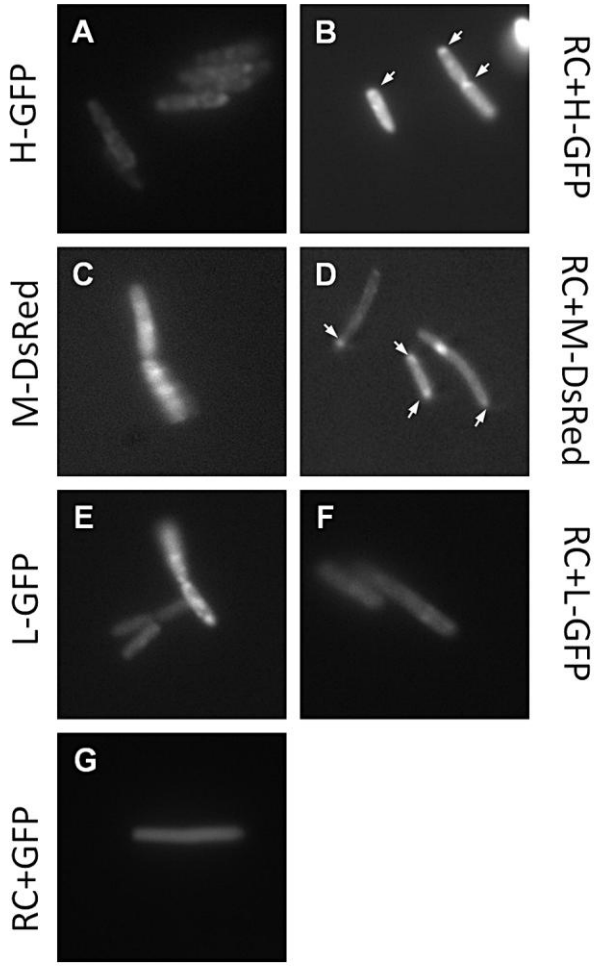

Figure 40.

Microscopy of RC protein fluorescent fusions in *E. coli* JM109. The left sets of panels (A, C and E) are showing the fluorescence of the cells containing a fusion of an individual RC subunit to an indicated fluorescent protein. The right sets of panels (B, D and F) are showing the same protein fusion when co-expressed with an unlabeled RC complex. Arrows indicate the shifting of fluorescence intensity towards the poles of the cells. Panel G shows the diffuse expression of GFP when co-expressed with the RC complex.

DISCUSSION

With the ultimate goal of engineering the light harvesting ability to nonphotosynthetic hosts, better understanding of the requirements for heterologous expression of the RC is required. This work was undertaken in order to gain those

insights and begin the initial steps of engineering light capture in a heterologous host. With the results presented here bringing us one step closer to that goal.

Several unique challenges presented themselves during this process, the first being the toxicity of RC membrane proteins when expressed in *E. coli*. The toxicity associated with their expression was obvious, as *E. coli* as cell growth slowed significantly and the final concentration of cells was reduced compared to non-RC containing controls. Many studies have published methods attempting to mediate the toxicity associated with membrane protein expression (reviewed in (23)). In our case, lowering the growth temperature to approximately 25°C as well as using a low copy (<10 copies/cell) pBBR backbone (28), appeared to help stabilize RC expression. Additionally, strain choice is known to matter in mediation of membrane protein toxicity (29, 30). Again, matching other reports, we saw a large variation of final RC yields dependent on the strain utilized, with *E. coli* JM109 cells proving least susceptible to the toxicity. Unfortunately, it is unclear at the present time which characteristic of *E. coli* JM109 cells makes them more suitable for RC expression, as it would be interesting to carry that trait into other protein expression strains.

In addition to optimizing the growth conditions and screening the strains used for expression, data on localization of the RC within the host cells gained through fluorescent fusions (Figure 39) may further aid in the choice of a future host. Interestingly, we saw the primarily polar localization of the RC when expressed in *E. coli*. One of the mechanisms for polar targeting of proteins in *E. coli* has been identified as the interaction of a target protein with cardiolipin (26). Furthermore, cardiolipin is known to interact

with the RC in *R. sphaeroides* (9, 10), and this interaction appears to be important in determining the localization of the complete RC complex at the poles. This suggests modulating cardiolipin levels in *E. coli* could help mediate some of the toxicity issues associated with RC expression by more evenly distributing the RC proteins across the entire cell membrane.

Not only did fluorescent protein fusions allowed us determine the localization of the RC, they helped to confirm *in vivo* that the reaction center subunits were assembling into a complete complex. Specifically, the change in localization of both H and M subunits from being diffuse in the cell, when expressed on their own, to aggregating near the poles upon coexpression with the RC complex (Figure 40). It is interesting that no significant change in localization of the L subunit was observed under the same conditions. While the L and the M subunits are highly similar, some differences do exist, and they could account for the different behavior observed. For example, even though both L and M subunit contain the cofactors needed for electron transfer from the special pair to the quinone, the L chain is favored in this process (6, 31). It is not clear if this favoritism and slight structural differences are the reason for different localization patterns we observed during our experiments. Knowing that the proteins composing the RC complex are able to co-localize when expressed in a heterologous host suggests that in the presence of the required cofactors the RC assembly could be active.

We initiated this work in an attempt to bring light capture to a nonphotosynthetic host. The results and new methods presented here bring us significantly closer to that goal. Once we are able to functionally express and reconstitute the RC with

bacteriochlorophylls within the cell, it should be possible to tie the energy captured by the RC into the cellular metabolism. Ultimately, we hope to be able to convert nonphotosynthetic hosts into light-capturing organisms by combining heterologous expression of the RC as described in this chapter with the complete bacteriochlorophyll biosynthetic pathway, as was described earlier in this thesis.

Conclusions and Future Directions.

This thesis focused on the central theme of introducing light-harvesting machinery into a nonphotosynthetic host in order to supplement the host's energy supplies. The additional energy could subsequently be utilized to offset the energy requirements from complex biosynthetic pathways. For a heterologous host to be able to capture light energy, reconstitution of a functional photosynthetic reaction center complex as well as the complete bacteriochlorophyll biosynthetic pathway is required. Neither of those tasks has been undertaken prior to the work described in this thesis.

While the work described here demonstrates significant progress towards our goals, significant research is still required to bring the ability to convert light into energy to a nonphotosynthetic host. In the case of the reaction center, we were able demonstrate stable expression of the complete complex in *E. coli*. Because the reaction center is essential for the process of light capture, this represents a major step towards our final goal and leads into the next crucial step, production of bacteriochlophylls by the host. The combination of the reaction centers with bacteriochlorophylls in a single host should form not only a stable, but also a functional light harvesting complex.

In order to bring the biosynthesis of bacteriochlorphylls to a heretologous host, in addition to the host's native heme biosynthetic pathway, expression of 16 other proteins is required. At the start of our work, the majority of the enzymes from both pathways were well characterized, with two exceptions being the 8-vinyl reductase (BciA) and Mg protorphyrin IX monomethylester cyclase (BchE). As presented in chapter 3, not every BciA homologue displayed sufficient *in vivo* activity when expressed in *E. coli* during

our attempts to extend the bacteriochlorophyll biosynthetic pathway. It is still unclear why the activities of the two BciA homologues differ so greatly and more work is needed to understand the differences between the two enzymes. A crystal structure of the proteins would be the most informative and effective way to begin to understand these differences. Small crystals were obtained during our work on research relating to chapter 3, though significant optimization is still required in order to solve the structure.

BchE, which was identified as the Mg protoporphyrin IX monomethylester cyclase over two decades ago, has never been purified or characterized in any meaningful way prior to this work. The mechanism of the cyclization reaction catalyzed by BchE remains largely a mystery. This makes the cyclization of MgPIX ME the only reaction from the bacteriochlorophyll biosynthesis about which no mechanistic details are know. While I was not able to demonstrate a complete catalytic turnover, I did develop the novel protocols for purifying and working with BchE. Furthermore, this work provides the first confirmations of the presence of the 4Fe-4S cluster, coupled with the interaction of BchE with SAM and predicted substrate significantly further our understanding of the enzyme. It is my hope that this data will be helpful to other researchers working on BchE and will help them to finally solve the mystery of BchE catalytic activity.

Ultimately, my work performed during my time as a graduate student, has brought us significantly closer to engineering light-energy conversion in a nonphotosynthetic host. Unfortunately, given the breadth of the project not everything is as complete as I would have wished it to be, but I hope my contributions will prove useful to future scientist working on similar problems.

References

Chapter 1:

1. Bryant, D.A. and N.U. Frigaard, *Prokaryotic photosynthesis and phototrophy illuminated.* Trends Microbiol, 2006. **14**(11): p. 488-96.
2. Liu, L.N. and S. Scheuring, *Investigation of photosynthetic membrane structure using atomic force microscopy.* Trends Plant Sci, 2013. **18**(5): p. 277-86.
3. Frigaard, N.U., et al., *Chlorobium tepidum: insights into the structure, physiology, and metabolism of a green sulfur bacterium derived from the complete genome sequence.* Photosynth Res, 2003. **78**(2): p. 93-117.
4. Tikkanen, M., et al., *Thylakoid protein phosphorylation in higher plant chloroplasts optimizes electron transfer under fluctuating light.* Plant Physiol, 2010. **152**(2): p. 723-35.
5. Hosler, J.P., S. Ferguson-Miller, and D.A. Mills, *Energy transduction: proton transfer through the respiratory complexes.* Annu Rev Biochem, 2006. **75**: p. 165-87.
6. Klamt, S., et al., *Modeling the electron transport chain of purple non-sulfur bacteria.* Mol Syst Biol, 2008. **4**: p. 156.
7. Gomez-Consarnau, L., et al., *Light stimulates growth of proteorhodopsin-containing marine Flavobacteria.* Nature, 2007. **445**(7124): p. 210-3.
8. Scholes, G.D., et al., *Lessons from nature about solar light harvesting.* Nat Chem, 2011. **3**(10): p. 763-74.
9. Eckhardt, U., B. Grimm, and S. Hortensteiner, *Recent advances in chlorophyll biosynthesis and breakdown in higher plants.* Plant Mol Biol, 2004. **56**(1): p. 1-14.
10. Scheer, H., *An Overview of Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, Functions and Applications*, in *Chlorophylls and Bacteriochlorophylls*, B. Grimm, et al., Editors. 2006, Springer Netherlands. p. 1-26.
11. Bollivar, D.W., *Recent advances in chlorophyll biosynthesis.* Photosynth Res, 2006. **90**(2): p. 173-94.
12. Gong, L. and S. Kaplan, *Translational control of puf operon expression in Rhodobacter sphaeroides 2.4.1.* Microbiology, 1996. **142 (Pt 8)**: p. 2057-69.
13. DeLong, E.F. and O. Beja, *The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times.* PLoS Biol, 2010. **8**(4): p. e1000359.
14. Lanyi, J.K., *Bacteriorhodopsin.* Annu Rev Physiol, 2004. **66**: p. 665-88.
15. Smith, S.O., *Structure and activation of the visual pigment rhodopsin.* Annu Rev Biophys, 2010. **39**: p. 309-28.
16. Oesterhelt, D. and W. Stoeckenius, *Functions of a new photoreceptor membrane.* Proc Natl Acad Sci U S A, 1973. **70**(10): p. 2853-7.
17. Beja, O., et al., *Bacterial rhodopsin: evidence for a new type of phototrophy in the sea.* Science, 2000. **289**(5486): p. 1902-6.
18. Marasco, E.K., K. Vay, and C. Schmidt-Dannert, *Identification of carotenoid cleavage dioxygenases from Nostoc sp. PCC 7120 with different cleavage activities.* J Biol Chem, 2006. **281**(42): p. 31583-93.
19. Walter, J.M., et al., *Light-powering Escherichia coli with proteorhodopsin.* Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2408-12.

20. Martinez, A., et al., *Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host.* Proc Natl Acad Sci U S A, 2007. **104**(13): p. 5590-5.
21. Kim, J.Y., et al., *Improved production of biohydrogen in light-powered Escherichia coli by co-expression of proteorhodopsin and heterologous hydrogenase.* Microb Cell Fact, 2012. **11**(1): p. 2.
22. Gonzalez, J.M., et al., *Genome analysis of the proteorhodopsin-containing marine bacterium Polaribacter sp. MED152 (Flavobacteria).* Proc Natl Acad Sci U S A, 2008. **105**(25): p. 8724-9.
23. Johnson, E.T., et al., *Enhancement of survival and electricity production in an engineered bacterium by light-driven proton pumping.* Appl Environ Microbiol, 2010. **76**(13): p. 4123-9.
24. Lavergne, J. and A. Vermeglio, *Functional coupling between reaction centers and cytochrome bc 1 complexes*. 2008. p. 509-536.
25. Goldsmith, J.O. and S.G. Boxer, *Rapid isolation of bacterial photosynthetic reaction centerss with an engineered poly-his tag.* Biochimica et biophysica acta, 1996. **1276**: p. 171-175.
26. Roszak, A.W., et al., *Crystal structure of the RC-LH1 core complex from Rhodopseudomonas palustris.* Science, 2003. **302**(5652): p. 1969-72.
27. Wraight, C. and M. Gunner, *The Acceptor Quinones of Purple Photosynthetic Bacteria — Structure and Spectroscopy*, in *The Purple Phototrophic Bacteria*. 2008. p. 379-405.
28. Handke, P., S.A. Lynch, and R.T. Gill, *Application and engineering of fatty acid biosynthesis in Escherichia coli for advanced fuels and chemicals.* Metab Eng, 2011. **13**(1): p. 28-37.
29. Zhang, F., S. Rodriguez, and J.D. Keasling, *Metabolic engineering of microbial pathways for advanced biofuels production.* Curr Opin Biotechnol, 2011. **22**(6): p. 775-83.
30. Fraser, N.J., H. Hashimoto, and R.J. Cogdell, *Carotenoids and bacterial photosynthesis: The story so far.* Photosynth Res, 2001. **70**(3): p. 249-56.
31. Maresca, J.A., J.E. Graham, and D.A. Bryant, *The biochemical basis for structural diversity in the carotenoids of chlorophototrophic bacteria.* Photosynth Res, 2008. **97**(2): p. 121-40.
32. Griffiths, M., et al., *Function of carotenoids in photosynthesis.* Nature, 1955. **176**(4495): p. 1211-5.
33. Monger, T.G., R.J. Cogdell, and W.W. Parson, *Triplet states of bacteriochlorophyll and carotenoids in chromatophores of photosynthetic bacteria.* Biochim Biophys Acta, 1976. **449**(1): p. 136-53.
34. Macpherson, A.N., et al., *Efficient energy transfer from the carotenoid S(2) state in a photosynthetic light-harvesting complex.* Biophys J, 2001. **80**(2): p. 923-30.
35. Lang, H.P. and C.N. Hunter, *The Relationship between Carotenoid Biosynthesis and the Assembly of the Light-Harvesting Lh2 Complex in Rhodobacter-Sphaeroides.* Biochemical Journal, 1994. **298**: p. 197-205.
36. Yeates, T.O., et al., *Structure of the reaction center from Rhodobacter sphaeroides R-26 and 2.4.1: protein-cofactor (bacteriochlorophyll, bacteriopheophytin, and carotenoid) interactions.* Proc Natl Acad Sci U S A, 1988. **85**(21): p. 7993-7.
37. Willows, R. and A. Kriegel, *Biosynthesis of bacteriochlorophylls in purple bacteria.* The purple phototrophic bacteria, 2008.

38. Johnson, E.T. and C. Schmidt-Dannert, *Characterization of three homologs of the large subunit of the magnesium chelatase from Chlorobaculum tepidum and interaction with the magnesium protoporphyrin IX methyltransferase.* J Biol Chem, 2008. **283**(41): p. 27776-84.
39. Bollivar, D.W., et al., *Heterologous expression of the bchM gene product from Rhodobacter capsulatus and demonstration that it encodes S-adenosyl-L-methionine:Mg-protoporphyrin IX methyltransferase.* J Bacteriol, 1994. **176**(17): p. 5290-6.
40. Ouchane, S., et al., *Aerobic and anaerobic Mg-protoporphyrin monomethyl ester cyclases in purple bacteria: a strategy adopted to bypass the repressive oxygen control system.* J Biol Chem, 2004. **279**(8): p. 6385-94.
41. Chew, A.G. and D.A. Bryant, *Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity.* Annu Rev Microbiol, 2007. **61**: p. 113-29.
42. Chew, A.G. and D.A. Bryant, *Characterization of a plant-like protochlorophyllide a divinyl reductase in green sulfur bacteria.* J Biol Chem, 2007. **282**(5): p. 2967-75.
43. Ito, H., et al., *Identification of a novel vinyl reductase gene essential for the biosynthesis of monovinyl chlorophyll in Synechocystis sp. PCC6803.* J Biol Chem, 2008. **283**(14): p. 9002-11.
44. Nagata, N., et al., *Identification of a vinyl reductase gene for chlorophyll synthesis in Arabidopsis thaliana and implications for the evolution of Prochlorococcus species.* Plant Cell, 2005. **17**(1): p. 233-40.
45. Bollivar, D.W., et al., *Directed mutational analysis of bacteriochlorophyll a biosynthesis in Rhodobacter capsulatus.* J Mol Biol, 1994. **237**(5): p. 622-40.
46. Rzeznicka, K., et al., *Xantha-l encodes a membrane subunit of the aerobic Mg-protoporphyrin IX monomethyl ester cyclase involved in chlorophyll biosynthesis.* Proc Natl Acad Sci U S A, 2005. **102**(16): p. 5886-91.
47. Tang, K.H., et al., *Role of the AcsF protein in Chloroflexus aurantiacus.* J Bacteriol, 2009. **191**(11): p. 3580-7.
48. Wong, Y.S. and P.A. Castelfranco, *Resolution and Reconstitution of Mg-Protoporphyrin IX Monomethyl Ester (Oxidative) Cyclase, the Enzyme System Responsible for the Formation of the Chlorophyll Isocyclic Ring.* Plant Physiol, 1984. **75**(3): p. 658-61.
49. Sawicki, A. and R.D. Willows, *BchJ and BchM interact in a 1 : 1 ratio with the magnesium chelatase BchH subunit of Rhodobacter capsulatus.* FEBS J, 2010. **277**(22): p. 4709-21.
50. Liu, Z. and D.A. Bryant, *Multiple types of 8-vinyl reductases for (bacterio)chlorophyll biosynthesis occur in many green sulfur bacteria.* J Bacteriol, 2011. **193**(18): p. 4996-8.

Chapter 2:

1. Hoff, A.J. and J. Deisenhofer, *Photophysics of photosynthesis. Structure and spectroscopy of reaction centers of purple bacteria.* Phys Rep, 1997. 287(1-2): p. 1-247.
2. Kobayashi, J., et al., *Efficient hydrogen production from acetate through isolated Rhodobacter sphaeroides.* J Biosci Bioeng, 2011. 112(6): p. 602-5.
3. Kien, N.B., et al., *Coenzyme Q10 production in a 150-l reactor by a mutant strain of Rhodobacter sphaeroides.* J Ind Microbiol Biotechnol, 2010. 37(5): p. 521-9.
4. Rajasekhar, N., C. Sasikala, and C.V. Ramana, *Photobiotransformation of indole to its value-added derivatives by Rhodobacter sphaeroides OU5.* J Ind Microbiol Biotechnol, 1998. 20(3-4): p. 177-179.
5. McKinlay, J.B. and C.S. Harwood, *Carbon dioxide fixation as a central redox cofactor recycling mechanism in bacteria.* Proc Natl Acad Sci U S A, 2010. 107(26): p. 11669-75.
6. Tabita, F.R., *Research on carbon dioxide fixation in photosynthetic microorganisms (1971-present).* Photosynth Res, 2004. 80(1-3): p. 315-332.
7. Zhao, H.P., et al., *Using a two-stage hydrogen-based membrane biofilm reactor (MBfR) to achieve complete perchlorate reduction in the presence of nitrate and sulfate.* Environ Sci Technol, 2013. 47(3): p. 1565-72.
8. Abo-Hashesh, M., N. Desaunay, and P.C. Hallenbeck, *High yield single stage conversion of glucose to hydrogen by photofermentation with continuous cultures of Rhodobacter capsulatus JP91.* Bioresour Technol, 2013. 128: p. 513-7.
9. Franchi, E., et al., *Metabolically engineered Rhodobacter sphaeroides RV strains for improved biohydrogen photoproduction combined with disposal of food wastes.* Mar Biotechnol (NY), 2004. 6(6): p. 552-65.
10. Kim, M.S., J.S. Baek, and J.K. Lee, *Comparison of H-2 accumulation by Rhodobacter sphaeroides KD131 and its uptake hydrogenase and PHB synthase deficient mutant.* Int J Hydrogen Energ, 2006. 31(1): p. 121-127.
11. Niederman, R.A., *Membrane development in purple photosynthetic bacteria in response to alterations in light intensity and oxygen tension.* Photosynth Res, 2013.
12. Adams, P.G. and C.N. Hunter, *Adaptation of intracytoplasmic membranes to altered light intensity in Rhodobacter sphaeroides.* Biochim Biophys Acta, 2012. 1817(9): p. 1616-27.
13. Chen, D.M., Y.B. Han, and Z.X. Gu, *Application of statistical methodology to the optimization of fermentative medium for carotenoids production by Rhodobacter sphaeroides.* Process Biochemistry, 2006. 41(8): p. 1773-1778.
14. Roy, A., et al., *Employing Rhodobacter sphaeroides to functionally express and purify human G protein-coupled receptors.* Biol Chem, 2008. 389(1): p. 69-78.
15. Hanson, D.K., D.L. Mielke, and P.D. Laible, *Harnessing Photosynthetic Bacteria for Membrane Protein Production*, in *Membrane Protein Crystallization*. 2009. p. 51-82.

16. Tehrani, A. and J.T. Beatty, *Effects of precise deletions in Rhodobacter sphaeroides reaction center genes on steady-state levels of reaction center proteins: a revised model for reaction center assembly.* Photosynth Res, 2004. 79(1): p. 101-108.
17. Jaschke, P.R., et al., *Modification of the genome of Rhodobacter sphaeroides and construction of synthetic operons.* Methods Enzymol, 2011. 497: p. 519-38.
18. Vasilyeva, L., et al., *Expression of luciferase gene under control of the puf promoter from Rhodobacter sphaeroides.* Appl Biochem Biotechnol, 1999. 77-9: p. 337-345.
19. Ind, A.C., et al., *Inducible-expression plasmid for Rhodobacter sphaeroides and Paracoccus denitrificans.* Appl Environ Microbiol, 2009. 75(20): p. 6613-5.
20. Gong, L. and S. Kaplan, *Translational control of puf operon expression in Rhodobacter sphaeroides 2.4.1.* Microbiology, 1996. 142 (Pt 8): p. 2057-69.
21. Shetty, R.P., D. Endy, and T.F. Knight, Jr., *Engineering BioBrick vectors from BioBrick parts.* J Biol Eng, 2008. 2: p. 5.
22. Canton, B., A. Labno, and D. Endy, *Refinement and standardization of synthetic biological parts and devices.* Nat Biotechnol, 2008. 26(7): p. 787-93.
23. Vick, J.E., et al., *Optimized compatible set of BioBrick vectors for metabolic pathway engineering.* Appl Microbiol Biotechnol, 2011. 92(6): p. 1275-86.
24. Strack, R.L., et al., *A noncytotoxic DsRed variant for whole-cell labeling.* Nat Methods, 2008. 5(11): p. 955-7.
25. Sistrom, W.R., *A requirement for sodium in the growth of Rhodopseudomonas spheroides.* J Gen Microbiol, 1960. 22: p. 778-85.
26. Saltikov, C.W. and D.K. Newman, *Genetic identification of a respiratory arsenate reductase.* Proc Natl Acad Sci U S A, 2003. 100(19): p. 10983-8.
27. Johnson, E.T., et al., *Enhancement of survival and electricity production in an engineered bacterium by light-driven proton pumping.* Appl Environ Microbiol, 2010. 76(13): p. 4123-9.
28. Kovach, M.E., et al., *Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes.* Gene, 1995. 166(1): p. 175-6.
29. Gomelsky, L., et al., *Hierarchical regulation of photosynthesis gene expression by the oxygen-responsive PrrBA and AppA-PpsR systems of Rhodobacter sphaeroides.* J Bacteriol, 2008. 190(24): p. 8106-14.
30. Beja, O., et al., *Bacterial rhodopsin: evidence for a new type of phototrophy in the sea.* Science, 2000. 289(5486): p. 1902-6.
31. Gourdon, P., et al., *Optimized in vitro and in vivo expression of proteorhodopsin: a seven-transmembrane proton pump.* Protein Expr Purif, 2008. 58(1): p. 103-13.
32. Stone, K.M., et al., *Structural insight into proteorhodopsin oligomers.* Biophys J, 2013. 104(2): p. 472-81.
33. Hu, Z., et al., *A powerful hybrid puc operon promoter tightly regulated by both IPTG and low oxygen level.* Biochemistry (Mosc), 2010. 75(4): p. 519-2.

34. Gomelsky, M. and S. Kaplan, *Molecular genetic analysis suggesting interactions between AppA and PpsR in regulation of photosynthesis gene expression in Rhodobacter sphaeroides 2.4.1.* J Bacteriol, 1997. 179(1): p. 128-134.
35. Yeliseev, A.A. and S. Kaplan, *A novel mechanism for the regulation of photosynthesis gene expression by the TspO outer membrane protein of Rhodobacter sphaeroides 2.4.1.* J Biol Chem, 1999. 274(30): p. 21234-21243.
36. Elsen, S., et al., *PpsR: a multifaceted regulator of photosynthesis gene expression in purple bacteria.* Mol Microbiol, 2005. 57(1): p. 17-26.
37. Lee, S.K., et al., *Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels.* Curr Opin Biotechnol, 2008. 19(6): p. 556-63.
38. Naylor, G.W., et al., *The photosynthesis gene cluster of Rhodobacter sphaeroides.* Photosynth Res, 1999. 62(2-3): p. 121-139.
39. Mank, N.N., et al., *Regulation of bacterial photosynthesis genes by the small noncoding RNA PcrZ.* Proc Natl Acad Sci U S A, 2012. 109(40): p. 16306-11.
40. Pfleger, B.F., et al., *Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes.* Nat Biotechnol, 2006. 24(8): p. 1027-32.
41. Salis, H.M., E.A. Mirsky, and C.A. Voigt, *Automated design of synthetic ribosome binding sites to control protein expression.* Nat Biotechnol, 2009. 27(10): p. 946-50.
42. Carrier, T.A. and J.D. Keasling, *Library of synthetic 5′ secondary structures to manipulate mRNA stability in Escherichia coli.* Biotechnol Prog, 1999. 15(1): p. 58-64.
43. Isaacs, F.J., et al., *Engineered riboregulators enable post-transcriptional control of gene expression.* Nat Biotechnol, 2004. 22(7): p. 841-7.

Chapter 3:

1. Rosgaard, L., et al., *Bioengineering of carbon fixation, biofuels, and biochemicals in cyanobacteria and plants.* J Biotechnol, 2012. **162**(1): p. 134-47.
2. Machado, I.M. and S. Atsumi, *Cyanobacterial biofuel production.* J Biotechnol, 2012. **162**(1): p. 50-6.
3. Work, V.H., et al., *Improving photosynthesis and metabolic networks for the competitive production of phototroph-derived biofuels.* Curr Opin Biotechnol, 2012. **23**(3): p. 290-7.
4. Maurino, V.G. and A.P. Weber, *Engineering photosynthesis in plants and synthetic microorganisms.* J Exp Bot, 2013. **64**(3): p. 743-51.
5. Johnson, E.T. and C. Schmidt-Dannert, *Light-energy conversion in engineered microorganisms.* Trends Biotechnol, 2008. **26**(12): p. 682-9.
6. Tikh, I. and C. Schmidt-Dannert, *Towards Engineered Light-Energy Conversion in Nonphotosynthetic Microorganisms*, in *Synthetic Biology: Tools and Applications*. 2013. p. 303-316.
7. Weeks, A.M. and M.C. Chang, *Constructing de novo biosynthetic pathways for chemical synthesis inside living cells.* Biochemistry, 2011. **50**(24): p. 5404-18.

8. Johnson, E.T., et al., *Enhancement of survival and electricity production in an engineered bacterium by light-driven proton pumping.* Appl Environ Microbiol, 2010. **76**(13): p. 4123-9.
9. Mijts, B.N., P.C. Lee, and C. Schmidt-Dannert, *Identification of a carotenoid oxygenase synthesizing acyclic xanthophylls: combinatorial biosynthesis and directed evolution.* Chem Biol, 2005. **12**(4): p. 453-60.
10. Mijts, B.N., P.C. Lee, and C. Schmidt-Dannert, *Engineering carotenoid biosynthetic pathways.* Methods Enzymol, 2004. **388**: p. 315-29.
11. Kwon, S.J., et al., *High-level production of porphyrins in metabolically engineered Escherichia coli: systematic extension of a pathway assembled from overexpressed genes involved in heme biosynthesis.* Appl Environ Microbiol, 2003. **69**(8): p. 4875-83.
12. Battersby, A.R., et al., *Biosynthesis of the pigments of life: formation of the macrocycle.* Nature, 1980. **285**(5759): p. 17-21.
13. Chew, A.G. and D.A. Bryant, *Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity.* Annu Rev Microbiol, 2007. **61**: p. 113-29.
14. Willows, R.D., *Biosynthesis of chlorophylls from protoporphyrin IX.* Nat Prod Rep, 2003. **20**(3): p. 327-41.
15. Johnson, E.T. and C. Schmidt-Dannert, *Characterization of three homologs of the large subunit of the magnesium chelatase from Chlorobaculum tepidum and interaction with the magnesium protoporphyrin IX methyltransferase.* J Biol Chem, 2008. **283**(41): p. 27776-84.
16. Bollivar, D.W., et al., *Heterologous expression of the bchM gene product from Rhodobacter capsulatus and demonstration that it encodes S-adenosyl-L-methionine:Mg-protoporphyrin IX methyltransferase.* J Bacteriol, 1994. **176**(17): p. 5290-6.
17. Ouchane, S., et al., *Aerobic and anaerobic Mg-protoporphyrin monomethyl ester cyclases in purple bacteria: a strategy adopted to bypass the repressive oxygen control system.* J Biol Chem, 2004. **279**(8): p. 6385-94.
18. Sawicki, A. and R.D. Willows, *BchJ and BchM interact in a 1 : 1 ratio with the magnesium chelatase BchH subunit of Rhodobacter capsulatus.* FEBS J, 2010. **277**(22): p. 4709-21.
19. Boldareva-Nuianzina, E.N., et al., *Distribution and origin of oxygen-dependent and oxygen-independent forms of Mg-protoporphyrin monomethylester cyclase among phototrophic proteobacteria.* Appl Environ Microbiol, 2013. **79**(8): p. 2596-604.
20. Tang, K.H., et al., *Role of the AcsF protein in Chloroflexus aurantiacus.* J Bacteriol, 2009. **191**(11): p. 3580-7.
21. Fujita, Y., et al., *Identification of a nifDK-like gene (ORF467) involved in the biosynthesis of chlorophyll in the cyanobacterium Plectonema boryanum.* Plant Cell Physiol, 1993. **34**(2): p. 305-14.
22. Schulz, R., et al., *Nucleotide sequence of a cDNA coding for the NADPH-protochlorophyllide oxidoreductase (PCR) of barley (Hordeum vulgare L.) and its expression in Escherichia coli.* Mol Gen Genet, 1989. **217**(2-3): p. 355-61.
23. Brocker, M.J., et al., *Crystal structure of the nitrogenase-like dark operative protochlorophyllide oxidoreductase catalytic complex (ChlN/ChlB)2.* J Biol Chem, 2010. **285**(35): p. 27336-45.
24. Moser, J., et al., *Structure of ADP-aluminium fluoride-stabilized protochlorophyllide oxidoreductase complex.* Proc Natl Acad Sci U S A, 2013. **110**(6): p. 2094-8.

25. Muraki, N., et al., *X-ray crystal structure of the light-independent protochlorophyllide reductase.* Nature, 2010. **465**(7294): p. 110-4.
26. Paddock, T., et al., *Arabidopsis light-dependent protochlorophyllide oxidoreductase A (PORA) is essential for normal plant growth and development.* Plant Mol Biol, 2012. **78**(4-5): p. 447-60.
27. Chew, A.G. and D.A. Bryant, *Characterization of a plant-like protochlorophyllide a divinyl reductase in green sulfur bacteria.* J Biol Chem, 2007. **282**(5): p. 2967-75.
28. Nagata, N., et al., *Identification of a vinyl reductase gene for chlorophyll synthesis in Arabidopsis thaliana and implications for the evolution of Prochlorococcus species.* Plant Cell, 2005. **17**(1): p. 233-40.
29. Parham, R. and C.A. Rebeiz, *Chloroplast biogenesis: [4-vinyl] chlorophyllide a reductase is a divinyl chlorophyllide a-specific, NADPH-dependent enzyme.* Biochemistry, 1992. **31**(36): p. 8460-4.
30. Canniffe, D.P., et al., *Identification of an 8-vinyl reductase involved in bacteriochlorophyll biosynthesis in Rhodobacter sphaeroides and evidence for the existence of a third distinct class of the enzyme.* Biochem J, 2013. **450**(2): p. 397-405.
31. Sytina, O.A., et al., *Ultrafast catalytic processes and conformational changes in the light-driven enzyme protochlorophyllide oxidoreductase (POR).* Biochem Soc Trans, 2009. **37**(Pt 2): p. 387-91.
32. Eckhardt, U., B. Grimm, and S. Hortensteiner, *Recent advances in chlorophyll biosynthesis and breakdown in higher plants.* Plant Mol Biol, 2004. **56**(1): p. 1-14.
33. Masuda, T., *Recent overview of the Mg branch of the tetrapyrrole biosynthesis leading to chlorophylls.* Photosynth Res, 2008. **96**(2): p. 121-43.
34. Shisler, K.A. and J.B. Broderick, *Emerging themes in radical SAM chemistry.* Curr Opin Struct Biol, 2012. **22**(6): p. 701-10.
35. Nagata, N., R. Tanaka, and A. Tanaka, *The major route for chlorophyll synthesis includes [3,8-divinyl]-chlorophyllide a reduction in Arabidopsis thaliana.* Plant Cell Physiol, 2007. **48**(12): p. 1803-8.
36. Wang, P., et al., *One divinyl reductase reduces the 8-vinyl groups in various intermediates of chlorophyll biosynthesis in a given higher plant species, but the isozyme differs between species.* Plant Physiol, 2013. **161**(1): p. 521-34.
37. Vick, J.E., et al., *Optimized compatible set of BioBrick vectors for metabolic pathway engineering.* Appl Microbiol Biotechnol, 2011. **92**(6): p. 1275-86.
38. Tamura, K., et al., *MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.* Mol Biol Evol, 2011. **28**(10): p. 2731-9.
39. Thompson, J.D., T.J. Gibson, and D.G. Higgins, *Multiple sequence alignment using ClustalW and ClustalX.* Curr Protoc Bioinformatics, 2002. **Chapter 2**: p. Unit 2 3.
40. Altschul, S.F., et al., *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.
41. Saitou, N. and M. Nei, *The neighbor-joining method: a new method for reconstructing phylogenetic trees.* Mol Biol Evol, 1987. **4**(4): p. 406-25.
42. Felsenstein, J., *Estimating effective population size from samples of sequences: a bootstrap Monte Carlo integration method.* Genet Res, 1992. **60**(3): p. 209-20.
43. Eswar, N., et al., *Comparative protein structure modeling using Modeller.* Curr Protoc Bioinformatics, 2006. **Chapter 5**: p. Unit 5 6.

44. Sreerama, N. and R.W. Woody, *Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set.* Anal Biochem, 2000. **287**(2): p. 252-60.
45. Mackenzie, C., et al., *Postgenomic adventures with Rhodobacter sphaeroides.* Annu Rev Microbiol, 2007. **61**: p. 283-307.
46. Mizoguchi, T., J. Harada, and H. Tamiaki, *Characterization of chlorophyll pigments in the mutant lacking 8-vinyl reductase of green photosynthetic bacterium Chlorobaculum tepidum.* Bioorg Med Chem, 2012. **20**(23): p. 6803-10.
47. Rescigno, M. and R.N. Perham, *Structure of the NADPH-binding motif of glutathione reductase: efficiency determined by evolution.* Biochemistry, 1994. **33**(19): p. 5721-7.
48. Pereira, P.J., et al., *Structure of human biliverdin IXβ reductase, an early fetal bilirubin IXβ producing enzyme.* Nat Struct Biol, 2001. **8**(3): p. 215-20.
49. Suzuki, J.Y. and C.E. Bauer, *Altered monovinyl and divinyl protochlorophyllide pools in bchJ mutants of Rhodobacter capsulatus. Possible monovinyl substrate discrimination of light-independent protochlorophyllide reductase.* J Biol Chem, 1995. **270**(8): p. 3732-40.
50. Kim, M.H., et al., *Structural insight into bioremediation of triphenylmethane dyes by Citrobacter sp. triphenylmethane reductase.* J Biol Chem, 2008. **283**(46): p. 31981-90.
51. Kolossov, V.L. and C.A. Rebeiz, *Chloroplast biogenesis 84: solubilization and partial purification of membrane-bound [4-vinyl]chlorophyllide a reductase from etiolated barley leaves.* Anal Biochem, 2001. **295**(2): p. 214-9.
52. Parham, R. and C.A. Rebeiz, *Chloroplast biogenesis 72: a [4-vinyl]chlorophyllide a reductase assay using divinyl chlorophyllide a as an exogenous substrate.* Anal Biochem, 1995. **231**(1): p. 164-9.
53. Bollivar, D.W., et al., *Directed mutational analysis of bacteriochlorophyll a biosynthesis in Rhodobacter capsulatus.* J Mol Biol, 1994. **237**(5): p. 622-40.
54. Griffiths, W.T., *Substrate-specificity studies on protochlorophyllide reductase in barley (Hordeum vulgare) etioplast membranes.* Biochem J, 1980. **186**(1): p. 267-78.
55. Townley, H.E., et al., *Protochlorophyllide oxidoreductase: a homology model examined by site-directed mutagenesis.* Proteins, 2001. **44**(3): p. 329-35.
56. Greenfield, N. and G.D. Fasman, *Computed circular dichroism spectra for the evaluation of protein conformation.* Biochemistry, 1969. **8**(10): p. 4108-16.
57. Huang, X., K. Nakanishi, and N. Berova, *Porphyrins and metalloporphyrins: versatile circular dichroic reporter groups for structural studies.* Chirality, 2000. **12**(4): p. 237-55.
58. Senge, M.O., *Stirring the porphyrin alphabet soup--functionalization reactions for porphyrins.* Chem Commun (Camb), 2011. **47**(7): p. 1943-60.
59. Liu, Z. and D.A. Bryant, *Multiple types of 8-vinyl reductases for (bacterio)chlorophyll biosynthesis occur in many green sulfur bacteria.* J Bacteriol, 2011. **193**(18): p. 4996-8.
60. Lopez-Gallego, F. and C. Schmidt-Dannert, *Multi-enzymatic synthesis.* Curr Opin Chem Biol, 2010. **14**(2): p. 174-83.
61. Hinchigeri, S.B., B. Hundle, and W.R. Richards, *Demonstration that the BchH protein of Rhodobacter capsulatus activates S-adenosyl-L-methionine:magnesium protoporphyrin IX methyltransferase.* FEBS Lett, 1997. **407**(3): p. 337-42.
62. Willows, R.D., et al., *Three separate proteins constitute the magnesium chelatase of Rhodobacter sphaeroides.* Eur J Biochem, 1996. **235**(1-2): p. 438-43.

63. Saunders, A.H., J.H. Golbeck, and D.A. Bryant, *Characterization of BciB, a ferredoxin-dependent 8-vinyl-protochlorophyllide reductase from the green sulfur bacterium Chloroherpeton thalassium.* Biochemistry, 2013.
64. Mijts, B.N. and C. Schmidt-Dannert, *Engineering of secondary metabolite pathways.* Curr Opin Biotechnol, 2003. **14**(6): p. 597-602.
65. Canniffe, D., et al., *Identification of an 8-vinyl reductase involved in bacteriochlorophyll biosynthesis in Rhodobacter sphaeroides and evidence for the existence of a third distinct class of the enzyme.* The Biochemical journal, 2013. **450**(2): p. 397-405.

Chapter 4:

1. Bryant DA, Frigaard NU. 2006. Prokaryotic photosynthesis and phototrophy illuminated. Trends Microbiol. 14:488-496.
2. Scheer H. 2006. An Overview of Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, Functions and Applications, p. 1-26. *In* Grimm B, Porra R, Rüdiger W, Scheer H (ed.), Chlorophylls and Bacteriochlorophylls, vol. 25. Springer Netherlands.
3. Chew AG, Bryant DA. 2007. Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity. Annu. Rev. Microbiol. 61:113-129.
4. Fraser NJ, Hashimoto H, Cogdell RJ. 2001. Carotenoids and bacterial photosynthesis: The story so far. Photosynth. Res. 70:249-256.
5. Yu J, Mathew S, Flavel BS, Johnston MR, Shapter JG. 2008. Ruthenium porphyrin functionalized single-walled carbon nanotube arrays--a step toward light harvesting antenna and multibit information storage. J. Am. Chem. Soc. 130:8788-8796.
6. Kralova J, Briza T, Moserova I, Dolensky B, Vasek P, Pouckova P, Kejik Z, Kaplanek R, Martasek P, Dvorak M, Kral V. 2008. Glycol porphyrin derivatives as potent photodynamic inducers of apoptosis in tumor cells. J. Med. Chem. 51:5964-5973.
7. Campbell WM, Jolley KW, Wagner P, Wagner K, Walsh PJ, Gordon KC, Schmidt-Mende L, Nazeeruddin MK, Wang Q, Gratzel M, Officer DL. 2007. Highly efficient porphyrin sensitizers for dye-sensitized solar cells. J Phys Chem C 111:11760-11762.
8. Song R, Kim YS, Sohn YS. 2002. Synthesis and selective tumor targeting properties of water soluble porphyrin-Pt(II) conjugates. J. Inorg. Biochem. 89:83-88.
9. Bollivar DW. 2006. Recent advances in chlorophyll biosynthesis. Photosynth. Res. 90:173-194.
10. Ouchane S, Steunou AS, Picaud M, Astier C. 2004. Aerobic and anaerobic Mg-protoporphyrin monomethyl ester cyclases in purple bacteria: a strategy adopted to bypass the repressive oxygen control system. J. Biol. Chem. 279:6385-6394.
11. Rzeznicka K, Walker CJ, Westergren T, Kannangara CG, von Wettstein D, Merchant S, Gough SP, Hansson M. 2005. Xantha-l encodes a membrane subunit of the aerobic Mg-protoporphyrin IX monomethyl ester cyclase involved in chlorophyll biosynthesis. Proc. Natl. Acad. Sci. U. S. A. 102:5886-5891.
12. Hunter CN, Coomber SA. 1988. Cloning and Oxygen-Regulated Expression of the Bacteriochlorophyll Biosynthesis Genes *bch* E, B, A and C of *Rhodobacter-Sphaeroides*. J. Gen. Microbiol. 134:1491-1497.

13. Bollivar DW, Beale SI. 1995. Formation of the Isocyclic Ring of Chlorophyll by Isolated *Chlamydomonas Reinhardtii* Chloroplasts. Photosynth. Res. 43:113-124.
14. Tottey S, Block MA, Allen M, Westergren T, Albrieux C, Scheller HV, Merchant S, Jensen PE. 2003. Arabidopsis CHL27, located in both envelope and thylakoid membranes, is required for the synthesis of protochlorophyllide. Proc. Natl. Acad. Sci. U. S. A. 100:16119-16124.
15. Minamizaki K, Mizoguchi T, Goto T, Tamiaki H, Fujita Y. 2008. Identification of two homologous genes, chlAI and chlAII, that are differentially involved in isocyclic ring formation of chlorophyll a in the cyanobacterium Synechocystis sp. PCC 6803. J. Biol. Chem. 283:2684-2692.
16. Bollivar DW, Beale SI. 1996. The Chlorophyll Biosynthetic Enzyme Mg-Protoporphyrin IX Monomethyl Ester (Oxidative) Cyclase (Characterization and Partial Purification from Chlamydomonas reinhardtii and Synechocystis sp. PCC 6803). Plant Physiol. 112:105-114.
17. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 39:D225-229.
18. Berthold DA, Stenmark P. 2003. Membrane-bound diiron carboxylate proteins. Annu. Rev. Plant Biol. 54:497-517.
19. Porra RJ, Urzinger M, Winkler J, Bubenzer C, Scheer H. 1998. Biosynthesis of the 3-Acetyl and 131-Oxo Groups of Bacteriochlorophyll a in the Facultative Aerobic Bacterium, *Rhodovulum sulfidophilum*. The presence of both oxygenase and hydratase pathways for isocyclic ring formation. Eur J of Biochem 257:185-191.
20. Bollivar DW, Suzuki JY, Beatty JT, Dobrowolski JM, Bauer CE. 1994. Directed mutational analysis of bacteriochlorophyll a biosynthesis in *Rhodobacter capsulatus*. J. Mol. Biol. 237:622-640.
21. Gough SP, Petersen BO, Duus JO. 2000. Anaerobic chlorophyll isocyclic ring formation in *Rhodobacter capsulatus* requires a cobalamin cofactor. Proc. Natl. Acad. Sci. U. S. A. 97:6908-6913.
22. Johnson ET, Schmidt-Dannert C. 2008. Characterization of three homologs of the large subunit of the magnesium chelatase from *Chlorobaculum tepidum* and interaction with the magnesium protoporphyrin IX methyltransferase. J. Biol. Chem. 283:27776-27784.
23. Kwon SJ, de Boer AL, Petri R, Schmidt-Dannert C. 2003. High-level production of porphyrins in metabolically engineered *Escherichia coli*: systematic extension of a pathway assembled from overexpressed genes involved in heme biosynthesis. Appl. Environ. Microbiol. 69:4875-4883.
24. Tehrani A, Thomas Beatty J. 2004. Effects of Precise Deletions in *Rhodobacter sphaeroides* Reaction Center Genes on Steady-state Levels of Reaction Center Proteins: A Revised Model for Reaction Center Assembly. Photosynth. Res. 79:101-108.
25. Vick JE, Johnson ET, Choudhary S, Bloch SE, Lopez-Gallego F, Srivastava P, Tikh IB, Wawrzyn GT, Schmidt-Dannert C. 2011. Optimized compatible set of BioBrick vectors for metabolic pathway engineering. Appl. Microbiol. Biotechnol. 92:1275-1286.

26. Pandey N, Sachan A, Chen Q, Ruebling-Jass K, Bhalla R, Panguluri KK, Rouviere PE, Cheng Q. 2013. Screening and identification of genetic loci involved in producing more/denser inclusion bodies in *Escherichia coli*. Microb Cell Fact 12:43.
27. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731-2739.
28. Cole C, Barber JD, Barton GJ. 2008. The Jpred 3 secondary structure prediction server. Nucleic Acids Res. 36:W197-201.
29. Saltikov CW, Newman DK. 2003. Genetic identification of a respiratory arsenate reductase. Proc. Natl. Acad. Sci. U. S. A. 100:10983-10988.
30. Sistrom WR. 1960. A requirement for sodium in the growth of *Rhodopseudomonas spheroides*. J. Gen. Microbiol. 22:778-785.
31. Chatterjee A, Li Y, Zhang Y, Grove TL, Lee M, Krebs C, Booker SJ, Begley TP, Ealick SE. 2008. Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. Nat. Chem. Biol. 4:758-765.
32. Tiwari MK, Lee JK, Moon HJ, Zhao H. 2011. Further biochemical studies on aminopyrrolnitrin oxygenase (PrnD). Bioorg. Med. Chem. Lett. 21:2873-2876.
33. Booker SJ. 2009. Anaerobic functionalization of unactivated C-H bonds. Curr. Opin. Chem. Biol. 13:58-73.
34. Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE. 2001. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. Nucleic Acids Res. 29:1097-1106.
35. Cheng S, Bobik TA. 2010. Characterization of the PduS cobalamin reductase of *Salmonella enterica* and its role in the Pdu microcompartment. J. Bacteriol. 192:5071-5080.
36. Bossemeyer D. 1994. The glycine-rich sequence of protein kinases: a multifunctional element. Trends Biochem. Sci. 19:201-205.
37. Berrisford JM, Sazanov LA. 2009. Structural basis for the mechanism of respiratory complex I. J. Biol. Chem. 284:29773-29783.
38. Bochkareva ES, Girshovich AS, Bibi E. 2002. Identification and characterization of the Escherichia coli stress protein UP12, a putative in vivo substrate of GroEL. Eur. J. Biochem. 269:3032-3040.
39. Brindley AA, Zajicek R, Warren MJ, Ferguson SJ, Rigby SE. 2010. NirJ, a radical SAM family member of the d1 heme biogenesis cluster. FEBS Lett. 584:2461-2466.
40. Fritsch J, Scheerer P, Frielingsdorf S, Kroschinsky S, Friedrich B, Lenz O, Spahn CM. 2011. The crystal structure of an oxygen-tolerant hydrogenase uncovers a novel iron-sulphur centre. Nature 479:249-252.
41. Nomata J, Ogawa T, Kitashima M, Inoue K, Fujita Y. 2008. NB-protein (BchN-BchB) of dark-operative protochlorophyllide reductase is the catalytic component containing oxygen-tolerant Fe-S clusters. FEBS Lett. 582:1346-1350.
42. Fluhe L, Knappe TA, Gattner MJ, Schafer A, Burghaus O, Linne U, Marahiel MA. 2012. The radical SAM enzyme AlbA catalyzes thioether bond formation in subtilosin A. Nat. Chem. Biol. 8:350-357.
43. Ollagnier-de Choudens S, Loiseau L, Sanakis Y, Barras F, Fontecave M. 2005. Quinolinate synthetase, an iron-sulfur enzyme in NAD biosynthesis. FEBS Lett. 579:3737-3743.

44. Wong. 1985. Resolution and reconstruction of Mg-protoporphryrin IX ME cyclase the enzyme system responsible for the formation of the chlorophyll isocyclic ring. Plant Physiol. 75.
45. Wong YS, Castelfranco PA, Goff DA, Smith KM. 1985. Intermediates in the formation of the chlorophyll isocyclic ring. Plant Physiol. 79:725-729.
46. Berthold DA, Stenmark P, Nordlund P. 2003. Screening for functional expression and overexpression of a family of diiron-containing interfacial membrane proteins using the univector recombination system. Protein Sci. 12:124-134.
47. Houry WA, Frishman D, Eckerskorn C, Lottspeich F, Hartl FU. 1999. Identification of in vivo substrates of the chaperonin GroEL. Nature 402:147-154.
48. Fujiwara K, Ishihama Y, Nakahigashi K, Soga T, Taguchi H. 2010. A systematic survey of in vivo obligate chaperonin-dependent substrates. EMBO J. 29:1552-1564.
49. Kerner MJ, Naylor DJ, Ishihama Y, Maier T, Chang HC, Stines AP, Georgopoulos C, Frishman D, Hayer-Hartl M, Mann M, Hartl FU. 2005. Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. Cell 122:209-220.
50. Layer G, Verfurth K, Mahlitz E, Jahn D. 2002. Oxygen-independent coproporphyrinogen-III oxidase HemN from *Escherichia coli*. J. Biol. Chem. 277:34136-34142.
51. Chan KK, Thompson S, O'Hagan D. 2013. The mechanisms of radical SAM/cobalamin methylations: an evolving working hypothesis. ChemBioChem 14:675-677.
52. Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE. 2001. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. Nucleic Acids Res. 29:1097-1106.
53. Davin LB, Lewis NG. 2000. Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis. Plant Physiol. 123:453-462.

Chapter 5:

1. Uchiyama, D., et al., *Reconstitution of bacterial photosynthetic unit in a lipid bilayer studied by single-molecule spectroscopy at 5 K.* Phys Chem Chem Phys, 2011. **13**(24): p. 11615-9.
2. Kalman, L., J.C. Williams, and J.P. Allen, *Comparison of bacterial reaction centers and photosystem II.* Photosynth Res, 2008. **98**(1-3): p. 643-55.
3. Söhlemann, P., C. Oeckl, and H. Michel, *Expression in Escherichia coli of the genes coding for reaction center subunits from Rhodobacter sphaeroides: wild-type proteins and fusion proteins containing one or four truncated domains from Staphylococcus aureus protein A at the carboxy-terminus.* Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression, 1991. **1089**(1): p. 103-112.
4. Axelrod, H.L., et al., *X-ray structure determination of the cytochrome c2: reaction center electron transfer complex from Rhodobacter sphaeroides.* J Mol Biol, 2002. **319**(2): p. 501-15.
5. Bryant, D.A. and N.U. Frigaard, *Prokaryotic photosynthesis and phototrophy illuminated.* Trends Microbiol, 2006. **14**(11): p. 488-96.

6. Okamura, M.Y., et al., *Proton and electron transfer in bacterial reaction centers.* Biochim Biophys Acta, 2000. **1458**(1): p. 148-63.
7. Paddock, M.L., G. Feher, and M.Y. Okamura, *Proton transfer pathways and mechanism in bacterial reaction centers.* FEBS Lett, 2003. **555**(1): p. 45-50.
8. Jones, M.R., *Lipids in photosynthetic reaction centres: structural roles and functional holes.* Prog Lipid Res, 2007. **46**(1): p. 56-87.
9. McAuley, K.E., et al., *Structural details of an interaction between cardiolipin and an integral membrane protein.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(26): p. 14706-14711.
10. Depalo, N., et al., *Enrichment of cardiolipin content throughout the purification procedure of photosystem II.* Bioelectrochemistry, 2004. **63**(1-2): p. 103-6.
11. Ventrella, A., et al., *Isolation and characterization of lipids strictly associated to PSII complexes: focus on cardiolipin structural and functional role.* Biochim Biophys Acta, 2007. **1768**(6): p. 1620-7.
12. Qian, P., et al., *Three-Dimensional Structure of the Rhodobacter sphaeroides RC-LH1-PufX Complex: Dimerization and Quinone Channels Promoted by PufX.* Biochemistry, 2013.
13. Saer, R.G., et al., *Role of Rhodobacter sphaeroides photosynthetic reaction center residue M214 in the composition, absorbance properties, and conformations of H(A) and B(A) cofactors.* Biochemistry, 2013. **52**(13): p. 2206-17.
14. Fyfe, P.K., et al., *Structural responses to cavity-creating mutations in an integral membrane protein.* Biochemistry, 2007. **46**(37): p. 10461-10472.
15. Chang, C.H., et al., *Structure of the membrane-bound protein photosynthetic reaction center from Rhodobacter sphaeroides.* Biochemistry, 1991. **30**(22): p. 5352-60.
16. Aird, A., et al., *Possible pathway for ubiquinone shuttling in Rhodospirillum rubrum revealed by molecular dynamics simulation.* Biophys J, 2007. **92**(1): p. 23-33.
17. Rinyu, L., et al., *Modulation of the free energy of the primary quinone acceptor (QA) in reaction centers from Rhodobacter sphaeroides: contributions from the protein and protein-lipid(cardiolipin) interactions.* Biochim Biophys Acta, 2004. **1655**(1-3): p. 93-101.
18. Klamt, S., et al., *Modeling the electron transport chain of purple non-sulfur bacteria.* Mol Syst Biol, 2008. **4**: p. 156.
19. Hosler, J.P., S. Ferguson-Miller, and D.A. Mills, *Energy transduction: proton transfer through the respiratory complexes.* Annu Rev Biochem, 2006. **75**: p. 165-87.
20. Lee, S.K., et al., *Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels.* Curr Opin Biotechnol, 2008. **19**(6): p. 556-63.
21. Tikh, I. and C. Schmidt-Dannert, *Towards Engineered Light-Energy Conversion in Nonphotosynthetic Microorganisms*, in *Synthetic Biology: Tools and Applications*. 2013. p. 303-316.
22. Goldsmith, J.O. and S.G. Boxer, *Rapid isolation of bacterial photosynthetic reaction centers with an engineered poly-histidine tag.* Biochimica Et Biophysica Acta-Bioenergetics, 1996. **1276**(3): p. 171-175.
23. Wagner, S., et al., *Rationalizing membrane protein overexpression.* Trends Biotechnol, 2006. **24**(8): p. 364-71.
24. Vick, J.E., et al., *Optimized compatible set of BioBrick vectors for metabolic pathway engineering.* Appl Microbiol Biotechnol, 2011. **92**(6): p. 1275-86.

25. Tehrani, A. and J. Thomas Beatty, *Effects of Precise Deletions in Rhodobacter sphaeroides Reaction Center Genes on Steady-state Levels of Reaction Center Proteins: A Revised Model for Reaction Center Assembly.* Photosynth Res, 2004. **79**(1): p. 101-8.
26. Romantsov, T., et al., *Cardiolipin promotes polar localization of osmosensory transporter ProP in Escherichia coli.* Mol Microbiol, 2007. **64**(6): p. 1455-65.
27. Mileykovskaya, E. and W. Dowhan, *Cardiolipin membrane domains in prokaryotes and eukaryotes.* Biochim Biophys Acta, 2009. **1788**(10): p. 2084-91.
28. Kovach, M.E., et al., *Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes.* Gene, 1995. **166**(1): p. 175-6.
29. Baneyx, F., *Recombinant protein expression in Escherichia coli.* Curr Opin Biotechnol, 1999. **10**(5): p. 411-21.
30. Drew, D., et al., *Optimization of membrane protein overexpression and purification using GFP fusions.* Nat Methods, 2006. **3**(4): p. 303-13.
31. Faries, K.M., et al., *High throughput engineering to revitalize a vestigial electron transfer pathway in bacterial photosynthetic reaction centers.* J Biol Chem, 2012. **287**(11): p. 8507-14.